ESTIMATING RELATIVE EFFICIENCY FROM PAIRED-COUNT DATA WITH OVER-DISPERSION, WITH APPLICATION TO FISHERY SURVEY CALIBRATION STUDIES

OSAMA M. BATAINEH

Estimating relative efficiency from paired-count data with over-dispersion, with application to fishery survey calibration studies

by

©Osama M. Bataineh

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirement for the Degree of Master of Science in Statistics

Department of Mathematics and Statistics
Memorial University of Newfoundland

February, 2008

St. John's Newfoundland Canada

Abstract

In many of the multi-species trawl surveys conducted by Fisheries and Oceans Canada, the survey vessel "Wilfred Templeman" (WT) may be replaced by the vessel "Alfred Needler" (AN). We examined paired-trawl experiments involving these two vessels to examine for differences in catchability. In particular, we examine for differences in catchability of the Witch flounder species.

The relative efficiency of the AN compared to the WT is defined as the ratio of the means from both trawl catches. Four models are investigated in this thesis, Conditional Poisson (i.e. Binomial), Mixed Binomial, Negative Binomial and Conditional Negative Binomial. When catch data are Poisson distributed ,the approach is clear and well-developed. However, over-dispersion creates problems, and over-dispersion is common in many types of data including fisheries data.

We dealt with the over-dispersion problem using the Negative Binomial distribution to model the paired-counts instead of the Poisson distribution. We develop Conditional Negative Binomial (Conditional NB) and Concentrated Negative Binomial (Concentrated NB) models for estimating relative efficiency. We compared estimates with those from the more commonly used approaches involving standard logistic regression and also a mixed binomial regression model. We found that the Conditional NB and mixed binomial models performed better.

Our results suggest that there were no significant differences in the relative catchability of the two vessels, based on the Mixed Binomial and Conditional NB models.

Acknowledgements

To give just an acknowledgment to Dr. Noel Cadigan would be an understatement, for his role and plethora of guidance given to me during my M.Sc. program and towards the completion of my thesis has been immense. His guidance has been as vital and as important, as one may imagine. With this, I would like to express my deep heart-felt gratitude to my supervisor, Dr. Noel Cadigan for his supervision, advice, encouragement, enduring patience and constant support. He has never been ceasing in his support in every possible way towards making me able, to accomplish my thesis. He has always provided clear explanations when I was (hopelessly) stuck, constantly driving the research with new and innovative research methodologies. It was undoubtedly a great privilege for me to work under his supervision for which I shall forever be grateful. I thank him whole heartily.

I want to sincerely acknowledge the financial support provided by the School of Graduate Studies, Department of Mathematics and Statistics, Dr. Noel Cadigan, Dr. Gary Sneddon and Dr. Paul Peng in the form of Graduate Assistantship & Teaching Assistantships.

I personally thank my very dearest friend Manish Madan, for the inspiring wisdom and encouragement that he has shared with me during his friendship with me. Words may not be just enough to thank him for what he has given me but let me assure, it will continue to hold a special place in my heart forever.

It is my great pleasure to thank all my friends and well-wishers who directly or indirectly encouraged and helped me in the M.Sc. program and contributed to this dissertation.

Last and most importantly, I am very much grateful to my parents, my sister and my younger brothers, for their strong belief in my capabilities and for all of their eternal love, support and encouragement during my stay far away from them.

To my parents whom I love the most ...

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background of the Problem

Assessment of fish stocks involves evaluating the status of a stock relative to its past. Stock indices are fundamental components of stock assessments. An index is a measurement that we expect is proportional to stock size. A random index R_y available for year y is related to stock size (S_y) via the model

$$E(R_y) = qS_y, \tag{1.1}$$

and S_y is treated as a fixed quantity to estimate. Often in stock assessment a population model is used to relate S_y with S_{y-1} and other quantities such as estimates of fishery catches and natural mortality. These population models contain parameters that need to be estimated, and stock indices are used for this purpose. The constant of proportionality, q, is usually referred to as the catchability index. Although we can not directly infer stock size from a time series of indices $R_1, \ldots, R_Y$, we can infer trends in stock size when q is the same each year. Note that q may be much different from one for many reasons; for example, the index may be based on a fishing gear that does not catch small fish, or the index may be based on measurements from only part of the stock area.

Stock size indices are often based on a survey in which randomly chosen sites are sampled for fish. Commonly used survey methods include random chosen sites. In our study, we focus on stratified random bottom trawl surveys such as those conducted off the east coast of Canada by Fisheries and Oceans Canada (e.g. Doubleday 1981). These are multi-species surveys that are used extensively in stock assessments. The information collected from these surveys is used for many other purposes as well, such as determining species at risk (e.g. Smedbol et. al. 2002) and evaluations related to closed areas. The survey observation is commonly referred to as a set (i.e. set the gear), or a tow when a trawl is used. The average survey catch can be taken as an index of stock size. If the same survey protocols are used from year to year then the catchability index should remain relatively constant.

When survey vessels are changed it is important to compare the efficiency and selectivity of the vessels for the species of interest. In bottom-trawl surveys conducted by Department of Fisheries and Oceans, the survey vessel "Wilfred Templeman" (WT) may be replaced by the vessel "Alfred Needler" (AN). These two vessels have potentially different catchabilities for some species. We examine paired-trawl experiments to estimate the relative difference between WT and AN catchabilities. This is often referred to as vessel calibration. We apply our methods to data for one specie, the Witch flounder (*Glyptocephalus cynoglussus*).

Sampling with nets or trawls remains a common technique for determining the relative abundance of aquatic organisms. In paired–trawl experiments two vessels are used to fish as close together as possible to minimize spatial heterogeneity between the stock densities the vessels encounter. The paired trawling is repeated at many different sites to cover a range of species, depths, and fish densities. Pelletier (1998) reviewed estimation methods used in many vessel calibration experiments. In the past, normal linear models for difference in log catches were used for analysis. This approach does not properly account for the stochastic nature of the data (counts). Benoit and Swain (2003) used a better approach by treating the catches from both vessels as Poisson or over-dispersed Poisson random variables (rv's), which are appropriate for count data, including zero counts. However, their approach was complicated because many fish density parameters usually had to be estimated. Additional details are provided later in this chapter. A similar approach was used by Pelletier (1998) with a mean-variance assumption that was the same as an over-dispersed Poisson distribution (e.g. Negative Binomial). To reduce the number of nuisance fish density parameters, Pelletier (1998) also assumed that fish densities were constant between paired tows, although this assumption will not be appropriate in a typical paired-trawl comparative fishing experiment.

In this thesis, stock densities are not assumed to be constant. We assume that stock densities are different at different paired-sites (i.e. between-pairs), which is usually the case in practice. There may be some spatial correlation between stock densities, but we do not try to utilize this in our analysis. Vessels are fished close together in a paired-trawl experiment, but it is not possible to ensure that exactly the same stock densities are fished by both vessels. A new development in this thesis involves methods to deal with random differences between within-pair stock densities. Note that we do not assume that differences in stock densities at different paired-tow locations are random. Catches from different paired-tow locations are assumed to be independent, but with location-specific means. Furthermore, we assume that the probability that a fish is captured is the same at each site and for all lengths but possibly different for each vessel (AN and WT). For some species the first two assumptions may not be appropriate. Length and location (e.g. depth) effects in capture probabilities are sometimes found. Our methods can be modified to account for such effects, although this is not pursued in this thesis. Our methods are directly applicable to species where length and depth effects in capture probabilities are unlikely.

The basic type of data collected in a comparative fishing experiment is paired count data, in which a sample of N pairs is obtained. The data is described in more details later in this chapter. The within-pair difference in log-means is assumed to be constant between pairs of trawls, but otherwise the means are assumed to be different.

This type of data is discussed in Section 4.5 of Cox and Snell (1989). These authors considered the case when the responses were Poisson rv's. In this thesis we consider the generalization that the responses are over-dispersed Poisson rv's, namely Negative Binomial.

The benchmark or simplest model for count data is the Poisson distribution. It is useful at the outset to review some fundamental properties that characterize the Poisson distribution. If the discrete rv Y is Poisson distributed with rate parameter $\lambda > 0$, then Y has density

$$P(Y = y) = \frac{e^{-\lambda}(\lambda)^y}{y!}, y = 0, 1, 2, ... \tag{1.2}$$

where $E(Y) = Var(Y) = \lambda$. This distribution has a single parameter λ and its k^{th} moment, $E(Y^k)$, may be derived by differentiating the moment generating function (mgf) k times

$$M_Y(t) \equiv E(e^{ty}) = \exp\left\{\lambda\left(e^t - 1\right)\right\}. \tag{1.3}$$

Equality of the mean and variance is referred to as the equi-dispersion property of the Poisson distribution. This property is frequently violated in real-life data. Overdispersion (or sometimes underdispersion) means the variance exceeds (or is less than) the mean. A key property of the Poisson distribution is additivity. That is if $Y_i \sim P(\lambda_i)$, i=1,2,... are independent rv's, and if $\sum \lambda_i < \infty$, then $\sum Y_i \sim P(\sum \lambda_i)$.

Let μ denote the common fish density encountered by each vessel at a tow station. If the number of fish entering a trawl as a Poisson process, and are caught independently with probability q, which is sometimes referred to as Poisson thinning (e.g. Grimmett and Stirzaker, 1992), then the catch will be Poisson distributed with mean $q \times \mu$. If q_1 and q_2 are the catchabilities of the WT and AN vessels, respectively, and Y_1 and Y_2 are their catches, then $Y_i \sim Poi(q_i\mu), i = 1, 2$. The relative efficiency of the WT compared to the AN is defined as the ratio of their means,

$$\rho = \frac{q_1}{q_2}. \tag{1.4}$$

Hence, the within-pair difference in log-means, or $log(\rho)$, is constant between tow sites, which is the same as the model considered in Cox and Snell (1989; Section 4.5). The fish densities (μ's) vary between paired-tow sites. If a total of N tow sites are sampled then there are $2N$ observations and $N+1$ parameters. The N density parameters, $\mu_1, ..., \mu_N$ are nuisance parameters that are not of direct interest but necessary to model the data. In our application $N = 57$. The only parameter of interest is the relative efficiency between the two vessels, ρ.

It is easier to use a conditional distribution that treats the sum of paired-trawl catches from both vessels as fixed. This eliminates the large number of fish density parameters, and the corresponding statistical likelihood function only involves ρ when catches are Poisson distributed. We show this below. The sum of catches from both trawls are treated like sample sizes. A thorough discussion on the roles of conditioning in statistical inference is provided by Reid (1995). Cox and Snell (1989) gave more detailed inferences about ρ for Poisson paired-count data using the conditional distribution of $Y_1 \mid (Y_1 + Y_2)$.

Millar (1992) advocated the conditional approach in closely related commercial fishing gear size selectivity studies when catches are Poisson distributed. In this case, selection curves can be fitted using the logistic regression generalized linear model (GLIM, see McCullagh and Nelder, 1989). The conditional approach has been used in paired-trawl calibration studies by Fanning (1985) and Lewy, Nielsen, and Hovgard (2004).

Benoît and Swain (2003) used an over-dispersion parameter to account for extra-Poisson variation, but they did not show clearly the reason behind the source of extra variation. In this thesis we explore in more detail the effect of over-dispersion on conditional inferences, but first we consider the simple situation when catches are Poisson rv's.

Let Y_1 and Y_2 be independent Poisson rv's for the number of fish caught at some tow station by the WT and AN, respectively. Let y_1 and y_2 denote observations of

Y_1 and Y_2. Also, let $E(Y_2) = q_2\mu = \lambda$ and $E(Y_1) = q_1\mu = \rho\lambda$. The conditional distribution of Y_1 given $Y_1 + Y_2$ is Binomial. This can be shown as follows.

$$\begin{aligned}
P(Y_1|Y_1+Y_2=n) &= \frac{P(Y_1=y_1)P(Y_2=y_2)}{P(Y_1+Y_2=n)} \\
&= \{\frac{e^{-\rho\lambda}(\rho\lambda)^{y_1}}{y_1!}\frac{e^{-\lambda}(\lambda)^{y_2}}{y_2!}\}/\{\frac{e^{-(\rho\lambda+\lambda)}(\rho\lambda+\lambda)^n}{n!}\} \\
&= \frac{n!}{y_1!y_2!}(\frac{\rho\lambda}{\rho\lambda+\lambda})^{y_1}(\frac{\lambda}{\rho\lambda+\lambda})^{y_2} \\
&= \frac{n!}{y_1!y_2!}(\frac{\rho}{1+\rho})^{y_1}(\frac{1}{1+\rho})^{y_2} \\
&= \frac{n!}{y_1!y_2!}p^{y_1}(1-p)^{n-y_1},
\end{aligned}$$

where $p = \rho/(1+\rho)$. This is the Binomial distribution with $n = y_1 + y_2$ and $p = \rho/(1+\rho)$. The only unknown parameter in this distribution is ρ. The λ parameter is eliminated in the conditional distribution.

Relative efficiency, ρ, is non-negative and constant for all sampling locations. To avoid complications due to boundary constraints it is better to estimate $log(\rho)$, and transform after for inferences about ρ. In this case, ρ can be defined as $\rho = \exp(\beta)$. This leads to

$$\frac{e^{\beta}}{1+e^{\beta}} = p(\beta),$$

where $p(\beta)$ is the canonical link function for the Binomial distribution GLIM (McCullagh and Nelder, 1989). If log relative efficiency is linearly related to covariates, such as length or depth (e.g. Benoit and Swain), then this produces a logistic regression model. This model is described in more detail in the next chapter.

In the Binomial distribution $E(Y_1) = np$ and $Var(Y_1) = np(1-p)$ or $Var(Y_1) = \phi np(1-p)$ for the over-dispersion case, where ϕ is an over-dispersion parameter. McCullagh and Nelder (1989) used a quasi-likelihood approach to deal with over-dispersion. Note that the over-dispersed Poisson approach to paired count data may leads to different statistical inferences (e.g. confidence intervals) than the over-dispersed Binomial approach. They are not exactly equivalent, whereas the Poisson and Binomial approaches without over-dispersion produce identical maximum likelihood estimates (MLE) and standard error's via the observed information matrix (Cox and Snell, 1989). The Binomial approach seems preferable, for reasons outlined in Cox and Snell (1989) and Reid (1995).

As mentioned previously, it is not possible to ensure that exactly the same stock densities (μ) are fished at each tow site. Let μ_1 and μ_2 denote the densities at each site fished by paired tows. Within-pair spatial heterogeneity in fish densities is a source of over-dispersion in paired-trawl comparative fishing studies (Lewy, Nielsen, and Hovgard, 2004). We account for this by using a mixture distribution, $\mu_i \sim Gamma(\mu, k)$, where μ is the gamma mean and μ^2/k is the variance. An over-dispersion parameter like ϕ may not be sufficient to account for these random effects. Note that the gamma mean μ is assumed to be different at different paired tow sites.

Conditional on μ_1 and μ_2, it is easy to show that $Y_1|Y$, where $Y = Y_1 + Y_2$, is still Binomially distributed with

$$\begin{aligned}
p = \frac{\rho}{1+\rho} &= \frac{\frac{q_1\mu_1}{q_2\mu_2}}{1+\frac{q_1\mu_1}{q_2\mu_2}} \\
p\left(1+\frac{q_1\mu_1}{q_2\mu_2}\right) &= \frac{q_1\mu_1}{q_2\mu_2} \\
p &= \rho\frac{\mu_1}{\mu_2}(1-p) \\
\frac{p}{1-p} &= \rho\frac{\mu_1}{\mu_2} \\
\log\left(\frac{p}{1-p}\right) &= \log\left(\rho\frac{\mu_1}{\mu_2}\right) \\
\log\left(\frac{p}{1-p}\right) &= \beta + log\left(\frac{\mu_1}{\mu_2}\right).
\end{aligned} \tag{1.5}$$

In Figure 1.1 we show 10000 realizations of the log-gamma ratio where each gamma has the same mean. It resembles a normal distribution which suggests that a reasonable approach to deal with within-pair variability in stock densities is a Binomial-normal mixture model, where $\delta = log(\mu_1/\mu_2)$ is assumed to be $N(0, \sigma^2)$. This is a common generalized linear mixed model (GLMM). GLMM's can structure multiple sources of variation, measured as covariates and unmeasured as random effects. They are described in more detail in the next chapter. In the GLMM , *p(β)* is defined as

$$\frac{e^{\beta+u}}{1+e^{\beta+u}} = p(\beta),$$

where u is a random effect and $u \sim N(0, \sigma^2)$. This is an approach we investigate to accommodate within-pair spatial variability in stock densities.

In the GLMM approach some information about the random effects (i.e. their variability) will also available in the marginal totals $Y_1, ..., Y_N$. It is not clear how efficient the GLMM approach may be. Also, proper selection of the random effects is required for valid point estimates and for correct standard errors when a nonlinear link function is used (e.g. Heagerty and Kurland 2001). However, it is relatively easy to show that the marginal (over gamma random effects) distribution of Y_i is Negative Binomial (NB). In the NB distribution, the probability of an event occurring is given by :

$$P(Y = y) = \frac{\Gamma(k+y)}{\Gamma(k)\Gamma(y+1)} \left(\frac{\mu}{\mu+k}\right)^y \left(\frac{k}{\mu+k}\right)^k, \tag{1.6}$$

where the mean is μ and the variance is $\mu(1+\mu/k)$. It is clear that the variance is greater than the mean. This form of the NB distribution, where $\mathrm{Var}(Y) \propto \mu^2$ when μ is large, is often considered to be a suitable distribution for modeling trawl catches (e.g. Gunderson, 1993), and in particular catches from pair-trawl fishing experiments (Pelletier, 1998). This is referred to as the NB2 model by Cameron and Trivedi (1998). A less common form is the NB1 distribution, with $\mathrm{Var}(Y) \propto \mu$. As $k \longrightarrow \infty$ then $Var(Y) \longrightarrow \mu$ (Poisson distribution) for the NB2 distribution. The parameter k measures the Poisson over-dispersion of the distribution. We also investigate the conditional NB2 distribution $Y_1|Y$ for inferences about ρ. This is an important new contribution of this thesis.

While the conditional Poisson distribution (i.e. Binomial) is very common, there appears to be little information published on the conditional NB distribution. An exception is Hausman, Hall, and Griliches (1984). They studied the conditional

NB1 distribution; however, Cadigan (in prep) has shown that this distribution is not suitable for trawl catches. The paper by James and Moser (1999) is another example of the suitability of the NB2 distribution for fishery catch data. In this thesis we study the conditional NB2 distribution, which we refer to as the conditional NB distribution for simplicity.

The NB is a discrete probability distribution that is often used for organism count data. Flexibility of the NB distribution to accommodate different values of k is an advantage when modeling frequency distributions. A characteristic of the NB distribution that lends itself particularly well to biological populations is that frequencies can decrease monotonically from a modal value, providing a highly skewed distribution. The Poisson distribution can also be skewed, but requires that the mean equals the variance, an assumption not required with the NB distribution.

The interpretation and derivation of the NB as a Poisson - gamma mixture is a result that can be algebraically derived in several different ways as in Greenwood and Yule (1920). Here, we approach the problem directly in terms of a mixture distribution. If a random variable Y (e.g. the number of fish caught in a tow) is conditionally distributed as Poisson(λ_i) given a fixed λ_i, then

$$f(y_i|\lambda_i) = \frac{\exp(-\lambda_i)\lambda_i^{y_i}}{y_i!}, \qquad y_i = 0, 1, \ldots.$$

and the mean parameter

$$E(y_i|\lambda_i) = \lambda_i. \tag{1.7}$$

Suppose the parameter λ_i is actually a random term with density function $g(\lambda_i)$. The marginal distribution of y is obtained by integrating out λ_i,

$$h(y_i) = \int f(y_i|\lambda_i)g(\lambda_i)d\lambda_i.$$

where $g(\lambda_i)$ is a mixing distribution. For specific choices of $f(.)$ and $g(.)$, for example Poisson and gamma densities respectively, the integral has an explicit solution. From here on the i subscript is omitted. Suppose that λ has a two-parameter gamma distribution $g(\lambda; k, \mu)$

$$g(\lambda; k, \mu) = \frac{\left(\frac{k}{\mu}\right)^k}{\Gamma(k)} \lambda^{k-1} e^{-\frac{k}{\mu}\lambda}, \quad k > 0, \ \mu > 0,$$

where $\mathrm{E}(\lambda) = \mu$ and $\mathrm{V}(\lambda) = \mu^2/k$.

The marginal distribution of y is given by

$$h(y|\mu, k) = \int \frac{\exp(-\lambda).\lambda^y}{y!} \frac{\left(\frac{k}{\mu}\right)^k}{\Gamma(k)} \lambda^{k-1} e^{-\frac{k}{\mu}\lambda} d\lambda. \tag{1.8}$$

Using the following definitions

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx, \quad a > 0$$

$$\Gamma(x - 1) = x!$$

$$\frac{\Gamma(a)}{b^a} = \int_0^\infty x^{a-1} e^{-bx} dx, \quad b > 0.$$

the integral in (1.8) can be re-written in the form

$$\begin{aligned} h(y|\mu, k) &= \frac{(k/\mu)^k}{\Gamma(k)\Gamma(y+1)} \int \exp\left\{-\lambda\left(1 + \frac{k}{\mu}\right)\right\} \lambda^{y+k-1} d\lambda \\ &= \frac{\left(\frac{k}{\mu}\right)^k \left(1 + \frac{k}{\mu}\right)^{-(k+y)} \Gamma(k+y)}{\Gamma(k)\Gamma(y+1)} \\ &= \frac{\Gamma(y+k)}{\Gamma(k)\Gamma(y+1)} \left(\frac{\mu}{k+\mu}\right)^y \left(\frac{k}{k+\mu}\right)^k. \end{aligned} \tag{1.9}$$

If k is an integer then

$$f(y) = \binom{y+k-1}{k-1} \left(\frac{\mu}{k+\mu}\right)^y \left(\frac{k}{k+\mu}\right)^k. \tag{1.10}$$

1.2 Objective of the Thesis

One of the main objectives of this thesis is to find a model that gives good estimates and confidence intervals for the relative efficiency. Conditional Poisson (i.e.

Bin), Overdispersed Binomial (OD Bin) and Mixed Binomial (GLMM) models will be studied for this purpose. We feel that trawl catches tend to follow the Negative Binomial distribution, so two more models will be studied, the Full Negative Binomial and the Conditional Negative Binomial (cond NB). In the full NB model, we simplify numerical techniques using a concentration approach, which is described in Chapter 2. Therefore, we refer to the full NB approach as the concentrated (conc) NB approach.

The second main objective of the thesis is to study the properties of the Conditional Negative Binomial model. If $Y_1 \sim NB(\mu_1, k)$ and $Y_2 \sim NB(\mu_2, k)$, then the conditional distribution of Y_1 given the total sum of Y_1 and Y_2 is :

$$P(Y_1 = y | Y_1 + Y_2 = n) = \frac{\frac{\Gamma(k+y)\Gamma(k+n-y)}{\Gamma(y+1)\Gamma(n-y+1)} \left(\frac{p}{q}\right)^y \left\{\frac{q(\mu_1+\mu_2)+k}{p(\mu_1+\mu_2)+k}\right\}^y}{\sum_{x=0}^{n} \frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)} \left(\frac{p}{q}\right)^x \left\{\frac{q(\mu_1+\mu_2)+k}{p(\mu_1+\mu_2)+k}\right\}^x}, \tag{1.11}$$

where $p = \frac{\mu_1}{\mu_1+\mu_2}$. This density function will be derived, studied, and analyzed in Chapter 2. It will be compared to previously mentioned distributions in terms of estimation and statistical inference (i.e. confidence intervals) for relative efficiency. The estimation of parameters for the five different models we investigate are presented in Chapter 2. The estimators will be applied to a case study involving Witch flounder (*Glyptocephalus cynoglussus*) in Chapter 3.

We use the well known maximum likelihood approach for estimation. We conduct a simulation study in the fourth chapter to examine the performance of the models under study. The five models (Bin, OD Bin, GLMM, Conc NB, Cond NB) will be compared in terms of bias and confidence intervals for relative efficiency. Our conclusions are outlined in Chapter 5.

Chapter 2

Model Developments

2.1 Introduction

In this chapter, methods are reviewed and/or developed for estimating differences in the relative efficiency of survey vessels, or more generally the ratio of means in paired count data. An important emphasis is on reliable confidence intervals. The common Binomial (Bin) and over-dispersed (OD Bin) models are reviewed, as well as an alternative Binomial mixed model (GLMM) for dealing with over-dispersion. Two approaches based on the NB distributional assumption for paired counts are also developed, which is new research. The general modeling approach used in this thesis is likelihood-based. This approach, and the associated maximum likelihood estimator, requires complete specification of the distribution of the responses. Statistical inference is usually performed under the assumption that the distribution is correctly specified, and this is the approach taken in this chapter.

The basic data structure and model assumption we considered is a sample of N pairs of counts, where the within-pair log difference in means is assumed to be constant across pairs, but otherwise the between-pairs means are different for each pair; that is, the data are

$$\begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \\ . & . \\ . & . \\ . & . \\ Y_{N1} & Y_{N2} \end{pmatrix},$$

where $E(Y_{ij}) = \mu_{ij}$,$i = 1, ..., N$, $j = 1, 2$ and $\log(\mu_{i1}/\mu_{i2}) = \beta$. This case is valid under distributions other than poisson random variable.

The results in this chapter are applied to a case study in Chapter 3, and further investigated in a simulation study in Chapter 4.

2.2 Likelihood Models

Likelihood models are based on specifying the joint density of the dependent variables. We assume that the random variable Y_i given the covariate vector $\mathbf{x}_i$ and parameter vector $\boldsymbol{\theta}$, is distributed with density f$(y_i|\mathbf{x}_i, \boldsymbol{\theta})$. The likelihood principle chooses as the estimator of $\boldsymbol{\theta}$ the value that maximizes the joint probability of observing the sample values $y_1, ..., y_n$. This probability, viewed as a function of parameters conditional on the data, is called the likelihood function and is denoted

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(y_i|\mathbf{x}_i, \boldsymbol{\theta}), \tag{2.1}$$

where independence over i is assumed. Maximizing the likelihood function is equivalent to maximizing the log-likelihood function

$$l(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f(y_i|\mathbf{x}_i, \boldsymbol{\theta}). \tag{2.2}$$

In this thesis, three different likelihood approaches will be used, namely, maximum likelihood, profile likelihood and conditional likelihood. The maximum likelihood approach will be used for estimation purposes in the next section, and the other two approaches will be used for the Negative Binomial model.

2.3 Poisson Model

2.3.1 Conditional Poisson

In this section, we present some details about a very important member of the family of generalized linear models, namely binomial logistic regression. As shown in Chapter 1, the binomial distribution can be derived from conditioning one of two Poisson random variables on the total sum of both of them. This takes us to logistic regression.

Logistic regression became a useful tool in the 1950s in applications of biostatistics. Consider first a regression structure in which the response is binary (0 or 1) if the endpoint of an experimental run is whether a fish is caught or not. It is reasonable to assume that the response is a Bernoulli random variable Y_i, where $E(Y_i) = p_i$, i=1,......,n. Here p_i is a probability in a Bernoulli process and $\mathrm{Var}(Y_i) = p_i(1 - p_i)$.

For grouped data, there are n_i experimental units at the ith data point, $i = 1, \ldots, n$. This is the case in our study where n_i fish are caught at each site i. Thus the model can be written as

$$E(Y_i) = n_i p_i, \quad i = 1, \ldots, N.$$

We assume that the p_i can be modeled as a function of a linear combination of known covariates; that is,

$$p_i = \frac{exp(\mathbf{x}_i'\boldsymbol{\beta})}{1 + exp(\mathbf{x}_i'\boldsymbol{\beta})}, \quad i = 1, 2, \ldots n$$

where $\mathbf{x}_i$ is a vector of predictor variables. In this case $\mathrm{Var}(Y_i) = n_i p_i(1 - p_i)$.

We use the maximum likelihood approach to estimate the parameter vector β. It is well known that the probability function for a single binomial random variable Y

indexed by n and p is given by $\binom{n}{y}p^y(1-p)^{n-y}$. Since $\binom{n}{y}$ for our situation does not involve β, it will be dropped, and thus the log likelihood for the logistic regression model is given by

$$\log\{L(\mathbf{p};\mathbf{y})\} = \sum_{i=1}^{n}\left\{y_i \log\left(\frac{p_i}{1-p_i}\right) + n_i \log(1-p_i)\right\}, \tag{2.3}$$

where $y_1, y_2, \ldots, y_n$ are observed values of independent binomial random variables. The term $\log\left(\frac{p_i}{1-p_i}\right)$ is called the **logit** and can be written as

$$\begin{aligned}\log\left(\frac{p_i}{1-p_i}\right) &= \mathbf{x}_i^{'}\boldsymbol{\beta}\\ &= \sum_{j=1}^{k} x_{ij}\beta_j, \quad i = 1, 2, \ldots, n.\end{aligned}$$

We assume that $n \geq k$. The loglikelihood can also be written as

$$\begin{aligned}\log\{L(\boldsymbol{\beta};\mathbf{y})\} &= \sum_{i=1}^{n}\sum_{j=1}^{k} y_i x_{ij}\beta_j - \sum_{i=1}^{n} n_i \log\left\{1 + exp(\sum_{j=1}^{k} x_{ij}\beta_j)\right\}\\ &= \boldsymbol{\beta}^{'}\mathbf{X}\mathbf{y} - \sum_{i=1}^{n} n_i \log\left\{1 + exp(\mathbf{x}_i^{'}\boldsymbol{\beta})\right\},\end{aligned} \tag{2.4}$$

where $\mathbf{X}$ is the traditional model matrix in linear regression and $\mathbf{y}$ is the response vector. The derivative with respect to $\boldsymbol{\beta}$ is

$$\frac{\partial \log\{L(\boldsymbol{\beta}_i;\mathbf{y})\}}{\partial\boldsymbol{\beta}} = \mathbf{X}^{'}\mathbf{y} - \sum_{i=1}^{n}\left\{\left(\frac{n_i}{1+e^{\mathbf{x}_i^{'}\boldsymbol{\beta}}}\right)e^{\mathbf{x}_i^{'}\boldsymbol{\beta}}\mathbf{x}_i\right\},$$

Since

$$\frac{e^{\mathbf{x}_i^{'}\boldsymbol{\beta}}}{1+e^{\mathbf{x}_i^{'}\boldsymbol{\beta}}} = p_i, \tag{2.5}$$

we have

$$\frac{\partial \log\{L(\boldsymbol{\beta};\mathbf{y})\}}{\partial\boldsymbol{\beta}} = \mathbf{X}^{'}\mathbf{y} - \sum_{i=1}^{n} n_i p_i \mathbf{x}_i.$$

Since the $n_i p_i$ represents the means of the binomial random variables, we can express the right-hand side above in matrix notation as $\mathbf{X}^{'}(\mathbf{y}-\boldsymbol{\mu})$, where

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}$$

and $\mu_i = n_i p_i$. The "μ" notation is motivated by the fact that at the ith data point the mean of the binomial distribution is given by $n_i p_i$. As a result, the maximum likelihood estimator (MLE) is the solution to the score equation

$$\mathbf{X}'(\mathbf{y} - \boldsymbol{\mu}) = \mathbf{0}. \tag{2.6}$$

2.3.2 Overdispersion

Overdispersion is an important concept that will interest us as we use logistic regression. When the response probabilities vary over groups of experimental units exposed to similar experimental conditions, some assumptions have to be made about the form of this variation. First we consider a general model, described by Williams (1982).

Suppose that the data consists of n proportions, y_i/n_i, i=1, 2, ..., n, and suppose that the corresponding response probability for the ith observation depends on k explanatory variables $X_1, X_2, ..., X_k$ through a linear logistic model. To introduce variability in the response probabilities, the actual response probability for the i observation, π_i, will be assumed to vary about a mean p_i. This response probability is therefore a random variable where E(π_i) = p_i. The variance of π_i must be zero when p_i is either zero or unity, and the simplest function for which this is true is

$$Var(\pi_i) = \tau p_i(1 - p_i). \tag{2.7}$$

where $\tau \geq 0$ is an unknown scale parameter. The quantity π_i is an unobservable random variable. The mean of Y_i, conditional on π_i, is given by

$$E(Y_i|\pi_i) = n_i\pi_i,$$

and the conditional variance of Y_i is

$$Var(Y_i|\pi_i) = n_i\pi_i(1 - \pi_i).$$

The unconditional mean and variance of Y_i is required to estimate the regression parameters. This leads us to investigate the effect of the assumption about the random variability in the response probabilities in equation (2.7) on $\mathrm{E}(Y_i)$ and $\mathrm{Var}(Y_i)$.

The unconditional expected value of a random variable Y can be obtained from the conditional expectation of Y given π using the equation

$$E(Y) = E\left\{E(Y|\pi)\right\}, \tag{2.8}$$

and the unconditional variance of Y is given by

$$Var(Y) = E\left\{Var(Y|\pi)\right\} + Var\left\{E(Y|\pi)\right\}. \tag{2.9}$$

Application of these two results gives

$$E(Y_i) = E\left\{E(Y_i|\pi_i)\right\} = E(n_i\pi_i) = n_iE(\pi_i) = n_ip_i,$$

and

$$Var(Y_i) = E\{Var(Y_i|\pi_i)\} + Var\{E(Y_i|\pi_i)\}.$$

Now,

$$\begin{aligned}
E\{Var(Y_i|\pi_i)\} &= E\{n_i\pi_i(1 - \pi_i)\} \\
&= n_i\{E(\pi_i) - E(\pi_i^2)\} \\
&= n_i\{E(\pi_i) - Var(\pi_i) - (E(\pi_i)^2\} \\
&= n_i\{p_i - \tau p_i(1 - p_i) - p_i^2\} \\
&= n_ip_i\left\{(1 - p_i)(1 - \tau)\right\}.
\end{aligned}$$

Also,

$$Var\{E(Y_i|\pi_i)\} = Var(n_i\pi_i) = n_i^2Var(\pi_i) = n_i^2\tau p_i(1 - p_i),$$

and so

$$Var(Y_i) = n_i p_i(1 - p_i)\{1 + (n_i - 1)\tau\}. \tag{2.10}$$

In the absence of random variation in the response probabilities, Y_i would have a binomial distribution, Bin(n_i, p_i), and in this case, Var(Y_i) $= n_i p_i(1 - p_i)$. This corresponds to the situation where $\tau = 0$ in equation (2.9) and leads to Var(Y_i)$=n_i p_i(1-p_i)$ in equation (2.10). If there is variation amongst the response probabilities, so that τ is greater than zero, the variance of Y_i will exceed $n_i p_i(1 - p_i)$, the variance under binomial sampling, by a factor of $\{1 + (n_i - 1)\tau\}$. Thus variation amongst the response probabilities causes the variance of the observed number of successes to be inflated resulting in overdispersion.

In the special case of ungrouped binary data $n_i = 1$, for all values of i, and the variance in equation (2.10) becomes $p_i(1-p_i)$, which is exactly the variance of a binary response variable. Consequently, binary data can provide no information about the parameter τ.

Suppose that evidence of overdispersion is found after fitting a linear logistic model to n observations of the form y_i/n_i, i =1, 2, ..., n. In order to model this overdispersion, the variance of Y_i will be taken to be $\phi_i n_i p_i(1 - p_i)$, where, from equation (2.10), $\phi_i = 1 + (n_i - 1)\tau$. This function includes an unknown parameter, τ, which will have to be estimated. $\tau > 0$ implies $\phi_i > 1$. If $\phi_i < 1$, we call the phenomenon underdispersion. However, this problem does not occur in practice as often as overdispersion.

We can use maximum likelihood to estimate the regression parameters and the dispersion parameter jointly. Williams (1982) shows how an estimate, $\hat{\tau}$, of the parameter τ can be found by equating the value of Pearson's χ-statistic for the model to its approximate expected value. The value of χ for a given model depends on the value of $\hat{\tau}$, and so this procedure is iterative. The estimate of the parameter τ will also depend on the actual explanatory variables in the fitted model.

The more common approach described in McCullagh and Nelder (1989) is based on the assumption that $Var(Y) = \phi n_i p_i(1 - p_i)$; that is, the Binomial overdispersion is constant between observations. In this case ϕ can be estimated more easily using the chi-square statistic, as outlined in McCullagh and Nelder (1989). This is the approach used in the *glm* function in R and this is the approach we use in this thesis. It is commonly applied in GLIM's, including those in Benoit and Swain (2003), and Lewy, Nielsen, and Hovgard (2004). McCullagh and Nelder (1989) state that overdispersion can arise in a number of ways. The simplest, and perhaps the most common mechanism, is clustering in the population, (e.g. Stigler 1986). Clusters usually vary in size, but McCullagh and Nelder (1989) assumed for simplicity that the cluster size, k, was fixed and that the n individuals were actually sampled from n/k clusters. In the ith cluster, the number of positive respondents, Z_i, is assumed to have a Binomial distribution with index k and parameter π_i, which varies from cluster to cluster. Thus the total number of positive respondents is

$$Y = Z_1 + Z_2 + \ldots + Z_{n/k}.$$

If we write $E(\pi_i) = \pi$ and Var(π_i)=$\tau^2\pi(1 - \pi)$, then using ([illegible]) and ([illegible]) it can be shown that the unconditional mean and variance of Y are

$$\begin{aligned} E(Y) &= np_i \\ Var(Y) &= np_i(1 - p_i)\{1 + (k - 1)\tau^2\} \\ &= \phi np_i(1 - p_i). \end{aligned}$$

Note that the overdispersion parameter $\phi = 1+(k\text{-}1)\tau^2$ depends on the cluster size and on the variability of π from cluster to cluster, but not on the sample size, n. This enables us to proceed as if the observations were binomially distributed and to estimate the dispersion parameter from the residuals. An estimate of ϕ can be based on the residual sum of squares appropriately weighted,

$$\hat{\phi} = \frac{1}{m - v} \sum_{i=1}^{m} \frac{(y_i - n_i\hat{p}_i)^2}{n_i\hat{p}_i(1 - \hat{p}_i)} = \chi^2/(m - v). \tag{2.11}$$

where m is the number of observations and v is the number of parameters. The last ϕ estimate in ( ) is a special case of

$$\hat{\phi} = \frac{1}{m-v}\sum_{i=1}^{m}\frac{\{y_i - E(Y_i)\}^2}{Var(Y_i)} = \chi^2/(m-v). \tag{2.12}$$

where χ^2 is the generalized Pearson statistic.

2.3.3 Mixed Binomial Model

A common feature of the models that have been described in the previous section is that their linear components contain terms known as fixed effects. Models for binary data may also incorporate what are known as random effects, and a model that contains a combination of fixed and random effects is known as a mixed model.

In this section, a mixed model for binomial data is described. The logistic regression model can be extended to include a single random effect. Suppose that there are n binomial observations of the form y_i/n_i, where y_i is the observed value of a binomial response, Y_i, associated with the ith proportion, $i = 1, 2, ..., n$. The binomial response variable will be assumed to depend on k explanatory variables, $X_1, X_2, ..., X_k$, which take the values $x_{1i}, x_{2i}, ..., x_{ki}$ for the ith observation, and on a random effect u_i.

If the corresponding response probability is p_i, the random variable Y_i has a binomial distribution conditional on u_i with parameters n_i and p_i. The dependence of the p_i on the explanatory variables and the random effect is then modeled by taking

$$\log(\frac{p_i}{1-p_i}) = \beta_0 + \beta_1 x_{1i} + ... + \beta_k x_{ki} + u_i, \tag{2.13}$$

where $\beta_0, \beta_1, ..., \beta_k$ are unknown fixed-effects parameters and u_i is a random effect.

The term u_i in ( ) is a realization of a random variable U_i, and we will assume that U_i has a normal distribution with zero mean and variance σ^2, that is $U_i \sim N(0, \sigma^2)$. The model in equation ( ) may also be written in the form

$$logit(p_i) = \beta_0 + \beta_1 x_{1i} + ... + \beta_k x_{ki} + \sigma_u z_i \tag{2.14}$$

where z_i is a realization of the standard normal random variable Z_i.

This model can be fitted through the method of maximum likelihood. Writing $\eta_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_k x_{ki}$ for the linear component of the model derived from the fixed effects, the model becomes

$$logit(p_i) = \eta_i + \sigma_u z_i,$$

and the likelihood of the ith of n observations is given by

$$\begin{aligned} L(\boldsymbol{\beta}, \sigma_u, z_i) &= \binom{n_i}{y_i} p_i^{y_i} (1 - p_i)^{n_i - y_i} \\ &= \binom{n_i}{y_i} \frac{\{exp(\eta_i + \sigma_u z_i)\}^{y_i}}{\{1 + exp(\eta_i + \sigma_u z_i)\}^{n_i}}. \end{aligned} \tag{2.15}$$

The likelihood of the n observations conditional on the random effects is $\prod_{i=1}^{n} L(\boldsymbol{\beta}, \sigma_u, z_i)$ which depends on the unknown parameters $\beta_0, \beta_1, ..., \beta_k$ and σ_u, and the unknown realizations of the random variables $Z_1, Z_2, ..., Z_n$ which have $N(0,1)$ distributions.

The standard method of handling a likelihood function that involves random variables that have a fully specified probability distribution is to integrate the likelihood function with respect to the distribution of these variables. After 'integrating out' the z, the resulting function is termed a marginal likelihood function, and depends only on $\beta_0, \beta_1, ..., \beta_k$ and σ_u. The maximum likelihood estimates of these parameters are then those values which maximize the marginal likelihood function given by

$$L(\boldsymbol{\beta}, \sigma_u) = \prod_{i=1}^{n} \int_{-\infty}^{\infty} \left[\binom{n_i}{y_i} \frac{\{exp(\eta_i + \sigma_u z_i)\}^{y_i}}{\{1 + exp(\eta_i + \sigma_u z_i)\}^{n_i}} \frac{exp(-z_i^2/2)}{\sqrt{2\pi}} \right] dz_i. \tag{2.16}$$

Usually, the logarithm of this marginal likelihood function is maximized using an optimization routine, but this is complicated by the fact that the integral can only be evaluated numerically. One way of carrying out this numerical integration is to use the Gauss-Hermite formula for numerical integration, or quadrature, according to which

$$\int_{-\infty}^{\infty} f(u) e^{-u^2} du \approx \sum_{r=1}^{m} c_r f(s_r). \tag{2.17}$$

where the values of c_r and s_r are given in standard tables, such as those of Abramowitz and Stegun (1972). The integral in equation ([illegible]) can then be expressed as a summation, and to the marginal likelihood is approximately

$$L(\boldsymbol{\beta}, \sigma_u) = \pi^{-n/2} \prod_{i=1}^{n} \left[\binom{n_i}{y_i} \sum_{r=1}^{m} c_r \frac{\{exp(\eta_i + \sigma_u s_r \sqrt{2})\}^{y_i}}{\{1 + exp(\eta_i + \sigma_u s_r \sqrt{2})\}^{n_i}} \right].$$

The values $\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_k$ and $\hat{\sigma}_u$, which maximize this expression, or its logarithm, can then be determined numerically. Standard errors of these parameter estimates are usually also available as a by-product of the optimization process. Standard errors are obtained from the inverse of H, where H is the hessian matrix of second derivatives of the negative log likelihood. This procedure can be implemented in R using the package *glmmML*.

This model can account for over-dispersion in the logit proportion of WT catch, caused by random differences in local within-pair stock densities fished by each vessel. If the local stock densities are viewed as independent and identically distributed (iid) from a gamma distribution then the logit proportion of catch depends on the log ratio of the two stock densities. A Normal random effect is a reasonable approximation to the distribution of the log ratio of gamma random variables (see Chapter 1).

Although in practice one is usually primarily interested in estimating the parameters in the marginal linear mixed-effects model (the fixed effects $\boldsymbol{\beta}$ and σ_u), it is often useful to calculate estimates for the random effects z_i as well.

One way of obtaining estimates of a random effect is through using an empirical Bayes procedure (e.g. Collett, 1991). In this approach, inference about a parameter θ is based on the following statement

$$p(\theta|\mathbf{y}) \propto L(\mathbf{y}|\theta)p(\theta), \tag{2.18}$$

where the constant of proportionality is

$$\int L(\mathbf{y}|\theta)p(\theta)d\theta. \tag{2.19}$$

which ensures that $p(\theta|\mathbf{y})$ integrates to unity.

In our study, the equivalent of the prior density, $p(\theta)$, is then the density of Z_i, the random variable corresponding to z_i, and the equivalent of $L(\mathbf{y}|\theta)$ is the likelihood function in equation ([illegible]) with $\boldsymbol{\beta}$ and σ_u replaced by their maximum likelihood estimates, $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}_u$. The posterior density of Z_i is then proportional to the product of $L(\hat{\boldsymbol{\beta}},\ \hat{\sigma}_u,\ z_i)$ and the density of Z_i, which from equation ([illegible]) is

$$\binom{n_i}{y_i}\left[\frac{\{exp(\hat{\eta}_i+\hat{\sigma}_u z_i)\}^{y_i}}{\{1+exp(\hat{\eta}_i+\hat{\sigma}_u z_i)\}^{n_i}}\right]\left\{\frac{exp(-z_i^2/2)}{\sqrt{2\pi}}\right\}. \tag{2.20}$$

where $\hat{\eta}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \ldots + \hat{\beta}_k x_{ki}$. An estimate of z_i is the value $\hat{z}_i$, which maximizes the logarithm of this function. On differentiating the logarithm of this expression with respect to z_i and equating the derivative to zero.

$$\begin{aligned}
\log\left\{L(\hat{\boldsymbol{\beta}},\hat{\sigma}_u,z_i)\right\} &= \log\left[\binom{n_i}{y_i}\left[\frac{\{exp(\hat{\eta}_i+\hat{\sigma}_u z_i)\}^{y_i}}{\{1+exp(\hat{\eta}_i+\hat{\sigma}_u z_i)\}^{n_i}}\right]\left\{\frac{exp(-z_i^2/2)}{\sqrt{2\pi}}\right\}\right] \\
&= \log\left\{\binom{n_i}{y_i}\right\} + y_i\left(\hat{\eta}_i+\hat{\sigma}_u z_i\right) - n_i\log\left\{1+exp\left(\hat{\eta}_i+\hat{\sigma}_u z_i\right)\right\} \\
&\quad - \frac{z_i^2}{2} - \log\left(\sqrt{2\pi}\right),
\end{aligned}$$

and

$$\frac{\partial\log\left\{L(\hat{\boldsymbol{\beta}},\hat{\sigma}_u,z_i)\right\}}{\partial z_i} = y_i\hat{\sigma}_u - n_i\frac{\hat{\sigma}_u exp\left(\hat{\eta}_i+\hat{\sigma}_u z_i\right)}{1+exp\left(\hat{\eta}_i+\hat{\sigma}_u z_i\right)} - z_i = 0$$

Finally, $\hat{z}_i$ can be found by numerically by solving for $\hat{z}_i$ the equation

$$\frac{n_i\hat{\sigma}_u exp(\hat{\eta}_i+\hat{\sigma}_u\hat{z}_i)}{1+exp(\hat{\eta}_i+\hat{\sigma}_u\hat{z}_i)} + \hat{z}_i = \hat{\sigma}_u y_i. \tag{2.21}$$

An estimate of of u_i is $\hat{u}_i = \hat{\sigma}_u\hat{z}_i$.

2.4 Negative Binomial Model

This section also deals with departures from the Poisson distribution. An alternative approach to deal with over-dispersion caused by within-pair local variation in stock densities is to work directly with the marginal distribution of paired-trawl catches. The mixed model approach and over-dispersed Binomial model approach discussed in the previous two sections may be inefficient because they do not utilize information about random effects provided by the total catches from both vessels, or more generally in the paired-sums. In the Poisson model Y_i has mean λ_i and variance λ_i. We now relax the variance assumption, because it is not suitable for many types of biological data. A common parametric model to account for overdispersion is the Negative Binomial (NB). This distribution arises when data are Poisson, but there is gamma-distributed unobserved individual heterogeneity such as when within-pair stock densities are iid gamma random variables.

The NB density function can be written as

$$f(y) = \frac{\Gamma(y+k)}{\Gamma(k)\Gamma(y+1)} \left(\frac{\mu}{k+\mu}\right)^y \left(\frac{k}{k+\mu}\right)^k,$$

where Γ is the gamma function. The parameter k is regarded as the overdispersion parameter. The mean and variance of Y are given by

$$E(Y) = \mu \tag{2.22}$$

$$Var(Y) = \mu\left(1 + \frac{\mu}{k}\right). \tag{2.23}$$

The leading motivation for considering parametric distributions other than the Poisson is that they have the potential to accommodate features of data that are inconsistent with Poisson assumptions. Some common departures from the Poisson model are as follows.

1. The failure of the mean equals variance restriction : Frequently the variance of data exceeds the mean, which is usually referred to as extra-Poisson variation or overdispersion relative to the Poisson model. If the variance is less than the mean, we have underdispersion.

2. The "excess zeros" or "zero inflation" problem : The observed data may show a higher relative frequency of zeros, or some other integer, than is consistent

with Poisson model. These cases are discussed in Mullahy (1986) and Lambert (1992). The higher relative frequency of zeros is a feature of all Poisson mixtures obtained by convolution.

3. Multimodality : Observed univariate count distributions are sometimes bimodal or multimodal. If this is also a feature of the conditional distribution of counts, perhaps because observations may be drawn from different populations, then extensions of the Poisson are desirable.

4. The failure of the conditional independence assumption : Event counts, especially if they are a time series, may be independent.

The last consideration has to do with the failure of the Poisson process assumption, whereas the first three are concessions to the characteristics of observed data. Note that in this thesis, the first consideration is applicable to paired-trawl data. This is the main reason we consider the NB distribution.

Figure [illegible] shows, as an example, the Poisson-gamma mixture (Negative Binomial), with mean 10 and $k = 5$ with the Poisson distribution with mean 10.

2.4.1 Concentrated Negative Binomial Model

In this section we consider direct estimation of relative efficiency (ρ) and the NB over-dispersion parameter k by the maximum likelihood method, but concentrating (or profiling) the stock density parameters $\mu_{.1}, \ldots, \mu_{.N}$ out from numerical estimation.

The full likelihood depends on a parameter vector $\boldsymbol{\mu}$ in addition to ρ and k, where

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_{1\cdot} \\ \mu_{2\cdot} \\ \vdots \\ \mu_{n\cdot} \end{pmatrix},$$

so the likelihood function is $L(\rho, k, \boldsymbol{\mu})$. The concentrated or profile likelihood eliminates $\boldsymbol{\mu}$ by obtaining the restricted MLE of $\hat{\boldsymbol{\mu}}$ for fixed ρ and k. Then

$$L_{Conc}(\rho, k) = L\left\{\rho, k, \hat{\boldsymbol{\mu}}(\rho, k)\right\}. \tag{2.24}$$

The profile likelihood is useful if $\boldsymbol{\mu}$ is a nuisance parameter. For example, our interest is in the parameter ρ, and k is important for confidence intervals. In such circumstances there is an advantage to profiling out $\boldsymbol{\mu}$, especially if $\boldsymbol{\mu}$ is of high dimension. More details can be found in Davidson and MacKinnon (1993).

Restricted MLE's of the NB μ parameters are derived as follows. Recall that our data are $Y_{i1} \sim \text{NB}(\mu_{i1}, k)$ and $Y_{i2} \sim \text{NB}(\mu_{i2}, k)$, i=1,...,n. Again let $\mu_{i.} = \mu_{i1} + \mu_{i2}$. The constant relative efficiency assumption implies that $\mu_{i1} = p\mu_{i.}$ and $\mu_{i2} = q\mu_{i.}$ where $q = 1-p$ and $logit(p) = \beta$. The conditional joint density for the 2N observations is

$$\begin{aligned}
&P(y_{11}, \ldots, y_{n1}, y_{12}, \ldots, y_{n2} | p, k, \mu_{1.}, \ldots, \mu_{N.}) \\
&= \prod_{i=1}^{n}\prod_{j=1}^{2} P(y_{ij} | p, k, \mu_{1.}, \ldots, \mu_{N.}) \\
&= \prod_{i=1}^{n}\prod_{j=1}^{2} \left\{ \frac{\Gamma(y_{ij}+k)}{\Gamma(k)\Gamma(y_{ij}+1)} \left(\frac{\mu_{ij}}{\mu_{ij}+k}\right)^{y_{ij}} \left(\frac{k}{k+\mu_{ij}}\right)^{k} \right\}, \quad i = 1, \ldots, N, j = 1, 2.
\end{aligned} \tag{2.25}$$

Note that in ( ) the μ_{ij} are defined in terms of p and $\mu_{1.}, \ldots, \mu_{N.}$ as shown above. The corresponding log-likelihood function is

$$\begin{aligned}
&l(y_{11}, \ldots, y_{n1}, y_{12}, \ldots, y_{n2} | p, k, \mu_{1.}, \ldots, \mu_{N.}) \\
&= \sum_{i=1}^{n} \left[\log\left\{ \frac{\Gamma(y_{i1}+k)}{\Gamma(k)\Gamma(y_{i1}+1)} \right\} + y_{i1}\log\left\{ \frac{p\mu_{i.}}{p\mu_{i.}+k} \right\} + k\log\left\{ \frac{k}{k+p\mu_{i.}} \right\} \right] \\
&+ \sum_{i=1}^{n} \left[\log\left\{ \frac{\Gamma(y_{i2}+k)}{\Gamma(k)\Gamma(y_{i2}+1)} \right\} + y_{i2}\log\left\{ \frac{q\mu_{i.}}{q\mu_{i.}+k} \right\} + k\log\left\{ \frac{k}{k+q\mu_{i.}} \right\} \right].
\end{aligned} \tag{2.26}$$

Differentiating with respect to $\mu_{i.}$ and setting to zero yields

$$\frac{\partial \log P(y_{11},\ldots,y_{n1},y_{12},\ldots,y_{n2}|p,k,\mu_1,\ldots,\mu_N)}{\partial \mu_{i.}}$$
$$= y_{i1}\left\{\frac{p^2\mu_{i.}+pk-p^2\mu_{i.}}{(p\mu_{i.}+k)^2}\right\}/\left(\frac{p\mu_{i.}}{p\mu_{i.}+k}\right)+k\left\{\frac{-kp}{(p\mu_{i.}+k)^2}\right\}/\left(\frac{k}{p\mu_{i.}+k}\right)$$
$$+\, y_{i2}\left\{\frac{q^2\mu_{i.}+qk-q^2\mu_{i.}}{(q\mu_{i.}+k)^2}\right\}/\left(\frac{q\mu_{i.}}{q\mu_{i.}+k}\right)+k\left\{\frac{-kq}{(q\mu_{i.}+k)^2}\right\}/\left(\frac{k}{q\mu_{i.}+k}\right)$$
$$= y_{i1}\frac{pk}{p\mu_{i.}(p\mu_{i.}+k)}-k\frac{p}{p\mu_{i.}+k}+y_{i2}\frac{qk}{q\mu_{i.}(q\mu_{i.}+k)}-k\frac{q}{q\mu_{i.}+k}$$
$$= \frac{y_{i1}k}{\mu_{i.}(p\mu_{i.}+k)}-\frac{kp}{p\mu_{i.}+k}+\frac{y_{i2}k}{\mu_{i.}(q\mu_{i.}+k)}-\frac{kq}{q\mu_{i.}+k}$$
$$= \frac{y_{i1}k-kp\mu_{i.}}{\mu_{i.}(p\mu_{i.}+k)}+\frac{y_{i2}k-kq\mu_{i.}}{\mu_{i.}(q\mu_{i.}+k)}.$$

Setting this equation equal to zero is equivalent to solving

$$= \frac{y_{i1}-p\mu_{i.}}{p\mu_{i.}+k}+\frac{y_{i2}-q\mu_{i.}}{q\mu_{i.}+k}=0,$$

which leads to

$$\frac{y_{i1}-p\mu_{i.}}{p\mu_{i.}+k} = -\frac{y_{i2}-q\mu_{i.}}{q\mu_{i.}+k}$$
$$\Rightarrow (y_{i1}-p\mu_{i.})(q\mu_{i.}+k) = -(p\mu_{i.}+k)(y_{i2}-q\mu_{i.})$$
$$\Rightarrow y_{i1}q\mu_{i.}+y_{i1}k-pq\mu_{i.}^2-p\mu_{i.}k = -(y_{i2}p\mu_{i.}+y_{i2}k-pq\mu_{i.}^2-q\mu_{i.}k)$$
$$\Rightarrow y_{i1}(1-p)\mu_{i.}+y_{i1}k-p(1-p)\mu_{i.}^2-p\mu_{i.}k = -\{y_{i2}p\mu_{i.}+y_{i2}k-p(1-p)\mu_{i.}^2 - (1-p)\mu_{i.}k\}$$
$$\Rightarrow y_{i1}\mu_{i.}-py_{i1}\mu_{i.}+y_{i1}k-p\mu_{i.}^2+p^2\mu_{i.}^2-p\mu_{i.}k = -y_{i2}p\mu_{i.}-y_{i2}k+p\mu_{i.}^2-p^2\mu_{i.}^2+k\mu_{i.} - pk\mu_{i.}.$$

This simplifies to

$$y_{i1}\mu_{i.}-py_{i1}\mu_{i.}+y_{i1}k+y_{i2}p\mu_{i.}+y_{i2}k-k\mu_{i.}-2p\mu_{i.}^2+2p^2\mu_{i.}^2 = 0$$

or

$$(-2p+2p^2)\mu_{i.}^2+(y_{i1}-py_{i1}+py_{i2}-k)\mu_{i.}+(y_{i1}+y_{i2})k = 0.$$

The last equation is a quadratic of the form $a\mu_{i.}^2 + b\mu_{i.} + c = 0$ where $a = -2p + 2p^2$, $b = y_{i1} - py_{i1} + py_{i2} - k$ and $c = (y_{i1} + y_{i2})k$. So the mle of $\mu_{i.}$ for fixed p and k is

$$\begin{aligned}
\hat{\mu}_{i.}(p, k) &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\
&= \frac{-(y_{i1} - py_{i1} + py_{i2} - k) \pm \sqrt{(y_{i1} - py_{i1} + py_{i2} - k)^2 - 4(-2p + 2p^2)\{(y_{i1} + y_{i2})k\}}}{2(-2p + 2p^2)}.
\end{aligned} \tag{2.27}$$

Substituting this back into ([illegible]) yields the concentrated log-likelihood function.

$$\begin{aligned}
&L_{Conc}(p, k) \\
&= \sum_{i=1}^{n} \left[\log\left\{ \frac{\Gamma(y_{i1} + k)}{\Gamma(k)\Gamma(y_{i1} + 1)} \right\} + y_{i1} \log\left\{ \frac{p\hat{\mu}_{i.}(p, k)}{p\hat{\mu}_{i.}(p, k) + k} \right\} + k \log\left\{ \frac{k}{k + p\hat{\mu}_{i.}(p, k)} \right\} \right] \\
&+ \sum_{i=1}^{n} \left[\log\left\{ \frac{\Gamma(y_{i2} + k)}{\Gamma(k)\Gamma(y_{i2} + 1)} \right\} + y_{i2} \log\left\{ \frac{(1 - p)\hat{\mu}_{i.}(p, k)}{(1 - p)\hat{\mu}_{i.}(p, k) + k} \right\} \right] \\
&+ \sum_{i=1}^{n} \left[k \log\left\{ \frac{k}{k + (1 - p)\hat{\mu}_{i.}(p, k)} \right\} \right].
\end{aligned} \tag{2.28}$$

Rather than estimate the $N + 2$ parameters in the full likelihood numerically, we estimate only 2 parameters in the concentrated likelihood given by ([illegible]). This greatly speeds estimation. The mle's for the $\mu_{i.}$ parameters are obtained using ([illegible]) and the mle's for p and k. The $\mu_{i.}$'s are nuisance parameters so we are not worried about finding standard errors for their estimates. Large sample standard errors for the mle's of p and k can be obtained from the inverse of the hessian matrix of the concentrated loglikelihood function evaluated at the mle values. We use numeric derivatives (hessian() function in R) to get the standard errors.

2.4.2 Conditional Negative Binomial Model

We have N pairs of observations, (Y_{11}, Y_{12}), ..., (Y_{n1}, Y_{n2}) that are NB distributed with mean and variances,

$$E(Y_{ij}) = \mu_{ij}, \quad Var(Y_{ij}) = \mu_{ij}\left(1 + \frac{\mu_{ij}}{k}\right), \quad i = 1, ..., N; \quad j = 1, 2.$$

All observations are independently distributed. We assume that the ratio of means is constant for all i, and we are interested in inferences about this ratio, $\rho = \mu_{i1}/\mu_{i2}$. This is a generalization of the Poisson model considered by Cox (1970, section 4.5.1). The generalization is to accommodate over-dispersion in the observations.

Define $\mu_{i.} = \mu_{i1} + \mu_{i2}$. We can specify the joint distribution of the observations in terms of N+2 parameters: $\rho, k, \mu_{1.}, ..., \mu_{N.}$. The parameter of interest is ρ, and the $\mu_{i.}'s$ and k are nuisance parameters that are not of direct interest but necessary in the distribution of (Y_{11}, Y_{12}), ..., (Y_{N1}, Y_{N2}). In Section 2.4.1 we presented a computationally more efficient method to estimate the $N + 2$ parameters directly from the $2N$ observations using a 2 dimensional numerical optimization; however, there is a problem with that approach.

It is well known that maximum likelihood estimates of variance parameters are seriously biased when there are many nuisance parameters. The mle of σ^2 in the normal linear regression model is a common example. If the number of parameters $p = N/2$ then the mle of σ^2 is biased by 50% of σ^2. This leads to poor statistical inferences unless some type of adjustment is made. Several other examples of problems with maximum likelihood estimation when the number of nuisance parameters is large are given in Barndorff-Neilsen and Cox (1994; Section 4.2). The approach we explore is conditioning, similar to what is standard when the observations are Poisson distributed (see Section 2.3.1). We explore the utility of using the conditional distribution of $(Y_{11}|Y_{1.}), ..., (Y_{N1}|Y_{N.})$ for inferences about ρ. Conditioning can be a useful approach when dealing with nuisance parameters.

Another argument for conditioning on $Y_{1.}, ..., Y_{N.}$ is that in the absence of any information about $\mu_{1.}, ..., \mu_{N.}$ the marginal distribution of $Y_{1.}, ..., Y_{N.}$ gives no additional information about ρ. The pair-totals really only define the precision of the data. Conditioning on their values make statistical inferences more relevant to the observed data, which is a position closer to the Bayesian philosophy. An extensive review of this subject is given by Reid (1995).

For the ith pair of observations (we drop the i subscript for now to simplify notation), the conditional distribution of $Y_1|Y_\cdot = n$ can be obtained from

$$P(Y_1 = y|Y_1 + Y_2 = n) = \frac{P(Y_1 = y, Y_2 = n - y)}{\sum_{x=0}^{n} P(Y_1 = x, Y_2 = n - x)}$$
$$= \frac{P(Y_1 = y)P(Y_2 = n - y)}{\sum_{x=0}^{n} P(Y_1 = x)P(Y_2 = n - x)}.$$

Note that Y_1 and Y_2 are independent so that

$$P(Y_1 = y_1, Y_2 = y_2) = P(Y_1 = y_1)P(Y_2 = y_2).$$

Recall that $p = \mu_1/(\mu_1 + \mu_2) = \rho/(1 + \rho)$ and q = 1-p and that the pdf of Y_i is

$$P(Y_i = y) = \frac{\Gamma(k + y)}{\Gamma(k)\Gamma(y + 1)} \left(\frac{\mu_i}{\mu_i + k}\right)^y \left(\frac{k}{\mu_i + k}\right)^k.$$

It follows that

$$P(Y_1 = y|Y_1 + Y_2 = n) = \frac{P(Y_1 = y)P(Y_2 = n - y)}{\sum_{x=0}^{n} P(Y_1 = x)P(Y_2 = n - x)},$$
$$= \frac{\left\{\frac{\Gamma(k+y)}{\Gamma(k)\Gamma(y+1)} \left(\frac{\mu_1}{\mu_1+k}\right)^y \left(\frac{k}{\mu_1+k}\right)^k\right\} \left\{\frac{\Gamma(k+n-y)}{\Gamma(k)\Gamma(n-y+1)} \left(\frac{\mu_2}{\mu_2+k}\right)^{n-y} \left(\frac{k}{\mu_2+k}\right)^k\right\}}{\sum_{x=0}^{n} \left[\left\{\frac{\Gamma(k+x)}{\Gamma(k)\Gamma(x+1)} \left(\frac{\mu_1}{\mu_1+k}\right)^x \left(\frac{k}{\mu_1+k}\right)^k\right\} \left\{\frac{\Gamma(k+n-x)}{\Gamma(k)\Gamma(n-x+1)} \left(\frac{\mu_2}{\mu_2+k}\right)^x \left(\frac{k}{\mu_2+k}\right)^k\right\}\right]}$$
$$= \frac{\frac{\Gamma(k+y)\Gamma(k+n-y)}{\Gamma(y+1)\Gamma(n-y+1)} \left(\frac{\mu_1}{\mu_2}\right)^y \left(\frac{1}{\mu_1+k}\right)^y \left(\frac{1}{\mu_2+k}\right)^{n-y}}{\sum_{x=0}^{n} \frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)} \left(\frac{\mu_1}{\mu_2}\right)^x \left(\frac{1}{\mu_1+k}\right)^x \left(\frac{1}{\mu_2+k}\right)^{n-x}}$$
$$= \frac{\frac{\Gamma(k+y)\Gamma(k+n-y)}{\Gamma(y+1)\Gamma(n-y+1)} \left(\frac{p\mu_\cdot}{q\mu_\cdot}\right)^y \left(\frac{\mu_2+k}{\mu_1+k}\right)^y \left(\frac{1}{\mu_2+k}\right)^n}{\sum_{x=0}^{n} \frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)} \left(\frac{p\mu_\cdot}{q\mu_\cdot}\right)^x \left(\frac{\mu_2+k}{\mu_1+k}\right)^x \left(\frac{1}{\mu_2+k}\right)^n}$$
$$= \frac{\frac{\Gamma(k+y)\Gamma(k+n-y)}{\Gamma(y+1)\Gamma(n-y+1)} \left(\frac{p}{q}\right)^y \left(\frac{q\mu_\cdot+k}{p\mu_\cdot+k}\right)^y}{\sum_{x=0}^{n} \frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)} \left(\frac{p}{q}\right)^x \left(\frac{q\mu_\cdot+k}{p\mu_\cdot+k}\right)^x}. \tag{2.29}$$

An algorithm to evaluate $p_n(x) = P(Y_1 = x|Y_\cdot = n)$ is given in Appendix A.

Note that we did not achieve a clean partition of the information about ρ, k and $\mu_\cdot$. We had hoped that the conditional distribution would only involve ρ and k and not $\mu_\cdot$; however, this is not the case. In the next section we explore how sensitive the conditional distribution is to values of $\mu_\cdot$.

2.4.3 Mean and Variance of Conditional NB

We could not find simple expressions for the conditional NB means and variances. They can be directly computed using

$$E(Y_1|Y_. = n) = \sum_{y_1=0}^{n} y_1 p_n(y_1); \quad E(Y_1^2|Y_. = n) = \sum_{y_1=0}^{n} y_1^2 p_n(y_1) \tag{2.30}$$

and

$$Var(Y_1|Y_. = n) = E(Y_1^2|Y_. = n) - \{E(Y_1|Y_. = n)\}^2. \tag{2.31}$$

In Figure [illegible] we plot the difference in the expected NB fraction $E(Y_1|Y_. = n)/n$ and the Binomial probability p. The results are identical when p=0.5 but can be substantially different otherwise, especially when k is small (i.e. the NB overdispersion is large). These results also suggest that $E(Y_1|Y_. = n)$ is sensitive to the value of $\mu_.$ which is a problem for estimation.

In Figure [illegible] we plot the variance of the conditional NB distribution. When n (total sum) is fixed the variance increases as k gets smaller (i.e. overdispersion gets larger). Also, the results are symmetric around p=0.5 for all k and n values. That is, the variance is the same for $p = p_o$ or $1 - p_o$. When $\mu_. >> n$ then the variance becomes more constant as a function of p.

To better understand the effect of $\mu_.$ on $E(Y_1|Y_. = n)$ we plot the expectation at various values of $\mu_.$ in Figures [illegible] for n = 1, 5, 25, 50, respectively. The heavy solid line is the NB conditional expectation. The dashed and dotted-dashed curves are Taylor's series approximations described later. The vertical dotted line in each panel denotes the value $\mu_. = n$. The shaded regions cover ±50% of n. The horizontal dotted line denotes the Binomial expectation np. Again, the expectation can be considerably different (and greater) than the Binomial result when k is small and $p < 0.5$.

The conditional NB expectation increases as $\mu_.$ increases or k decreases. The mean of the conditional NB distribution is a finite sum; hence, limiting moments can be found by evaluating the moments for the limiting probability density function. Let $p_{lim}(x) \equiv \lim_{\mu_. \to \infty} P(Y_1 = x | Y_. = n)$,

$$
\begin{aligned}
p_{lim}(x) &= \lim_{\mu_. \to \infty} \frac{\frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)} \left(\frac{p}{q}\right)^x \left(\frac{q\mu_.+k}{p\mu_.+k}\right)^x}{\sum_{x=0}^{n} \frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)} \left(\frac{p}{q}\right)^x \left(\frac{q\mu_.+k}{p\mu_.+k}\right)^x} \\
&= \frac{\frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)} \left(\frac{p}{q}\right)^x \left(\frac{q}{p}\right)^x}{\sum_{x=0}^{n} \frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)} \left(\frac{p}{q}\right)^x \left(\frac{q}{p}\right)^x} \\
&= \frac{\frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)}}{\sum_{x=0}^{n} \frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)}} \\
&= \frac{\lambda(x)}{\sum_{x=0}^{n} \lambda(x)},
\end{aligned}
$$

where

$$\lambda(x) = \frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)}.$$

The ratio of $p_{lim}(x+1)/p_{lim}(x)$ can be helpful in finding the first two moments of the limiting distribution. To simplify notation let X denote a random variable with the limiting distribution of $Y_1 | Y_. = n$.

$$
\begin{aligned}
&\frac{p_{lim}(x+1)}{p_{lim}(x)} \\
&= \frac{\lambda(x+1)}{\lambda(x)} \\
&= \frac{\Gamma(k+x+1)\Gamma(k+n-x-1)}{\Gamma(x+2)\Gamma(n-x)} / \frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)} \\
&= \frac{(x+k)(n-x)}{(x+1)(n-x+k-1)},
\end{aligned}
\tag{2.32}
$$

and this Equation (2.32) implies that

$$p_{lim}(x+1)\{(n+k)(x+1) - (x+1)^2\} = (k+x)(n-x)p_{lim}(x). \tag{2.33}$$

The sum of both sides of (2.33) for x = 0, ..., n - 1 is

$$(n+k)\sum_{x=0}^{n-1}(x+1)p_{lim}(x+1) - \sum_{x=0}^{n-1}(x+1)^2 p_{lim}(x+1)$$

$$= nk\sum_{x=0}^{n-1} p_{lim}(x) + (n-k)\sum_{x=0}^{n-1} xp_{lim}(x) - \sum_{x=0}^{n-1} x^2 p_{lim}(x).$$

Recall that $X \equiv Y_1|Y_. = n$, so

$$\sum_{x=0}^{n-1}(x+1)p_{lim}(x+1) = \sum_{x=1}^{n} xp_{\text{lim}}(x) = \sum_{x=0}^{n} xp_{\text{lim}}(x) = E(X),$$

and

$$\sum_{x=0}^{n-1}(x+1)^2 p_{lim}(x+1) = E(X^2).$$

Hence,

$$\begin{aligned}
(n+k)E(X) - E(X^2) &= nk\{1 - p_{lim}(n)\} + (n-k)\{E(X) - np_{lim}(n)\} - E(X^2) \\
&\quad + n^2 p_{lim}(n), \\
\Longrightarrow (n+k)E(X) &= (n-k)E(X) - n(n-k)p_{lim}(n) + n^2 p_{lim}(n) \\
&\quad + nk\{1 - p_{lim}(n)\} \\
2kE(X) &= -n(n-k)p_{lim}(n) + n^2 p_{lim}(n) + nk\{1 - p_{lim}(n)\} \\
E(X) &= -\frac{n(n-k)p_{lim}(n)}{2k} + \frac{n^2 p_{lim}(n)}{2k} + \frac{nk\{1 - p_{lim}(n)\}}{2k} \\
&= -\frac{n^2 p_{lim}(n)}{2k} + \frac{np_{lim}(n)}{2} + \frac{n^2 p_{lim}(n)}{2k} + \frac{n}{2} - \frac{np_{lim}(n)}{2} \\
&= \frac{n}{2}.
\end{aligned}$$

We can also show that

$$\lim_{k\to 0} E(Y_1|Y_. = n; \mu_., k) = \frac{n}{2}.$$

This suggests that when $\mu_.$ is large or when the NB overdispersion is large (i.e. k is small) then the data are less informative about ρ. This is because as $\mu_.$ or the overdispersion increases then $E(Y_1|Y_. = n)$ deviates more from np, where $p = \rho/(1+\rho)$, and eventually this expectation equals n/2 regardless of the value of ρ. For finite values of $\mu_.$ and k, direct estimation of ρ from $\widetilde{p} = y_1/y_.$ will be biased towards one (i.e. equal catchability). For example, if $n = \mu_. = 25$, $p = 0.25$, and $k = 0.5$ then $E(Y_1|Y_. = n) \doteq 8.5$ (see Figure). If we were to observe this value then $\tilde{p} = 8.5/25 = 0.34$ and $\tilde{\rho} = 0.34/0.66 = 0.52$, whereas the true value is $\rho = 0.25/0.75 = 0.33$. This is a bias in the estimate of ρ towards one. This is similar to the attenuation bias resulting from measurement errors in covariates in linear regression (e.g. Stefanski, 2000). Overdispersion will tend to "mask" differences in catchability and reduce power.

The conditional variance is shown in Figure . The different line types are for different values of k, which are shown in the top left-hand panel. The horizontal dotted line denotes the Binomial variability.

The Conditional NB variance increases as $\mu_.$ increases or k decreases. To help derive the limiting variance as $\mu_. \to \infty$ we multiply ( ) by $x + 1$ and sum it for $x = 0, \ldots, n - 1$,

$$
\begin{aligned}
&(n+k)\sum_{x=0}^{n-1}(x+1)^2 p_{lim}(x+1) - \sum_{x=0}^{n-1}(x+1)^3 p_{lim}(x+1) \\
&= nk\sum_{x=0}^{n-1}(x+1)p_{lim}(x) + (n-k)\sum_{x=0}^{n-1}x(x+1)p_{lim}(x) \\
&- \sum_{x=0}^{n-1}x^2(x+1)p_{lim}(x).
\end{aligned}
$$

It follows that

$$
\begin{aligned}
(n+k)E(X^2) - E(X^3) &= nk\{1 - p_{lim}(n)\} + (nk + n - k)\{n/2 - np_{lim}(n)\} \\
&\quad + (n-k-1)E(X^2) - n^2(n-k-1)p_{lim}(n) - E(X^3) \\
&\quad + n^3 p_{lim}(n). \\
(2k+1)E(X^2) &= nk - nkp_{lim}(n) + (nk+n-k)(n/2) - n^2 kp_{lim}(n) \\
&\quad - n^2 p_{lim}(n) + nkp_{lim}(n) - n^3 p_{lim}(n) + n^2 kp_{lim}(n) \\
&\quad + n^2 p_{lim}(n) + n^3 p_{lim}(n). \\
(2k+1)E(X^2) &= nk + (nk+n-k)(n/2)
\end{aligned}
$$

$$
\begin{aligned}
(2k+1)E(X^2) &= nk + n^2k/2 + n^2/2 - nk/2. \\
E(X^2) &= \frac{nk + n^2k/2 + n^2/2 - nk/2.}{2k+1} \\
E(X^2) &= \frac{nk}{2(2k+1)} + \frac{n^2k}{2(2k+1)} + \frac{n^2}{2(2k+1)} \\
&= \frac{nk + n^2k + n^2}{2(2k+1)}.
\end{aligned}
$$

Subtracting $E(X)^2$ from $E(X^2)$ gives

$$
\begin{aligned}
E(X^2) - E(X)^2 &= \frac{nk + n^2k + n^2}{2(2k+1)} - \left(\frac{n}{2}\right)^2 \\
&= \frac{nk + n^2k + n^2}{2(2k+1)} - \frac{2(2k+1)}{2(2k+1)}\left(\frac{n}{2}\right)^2 \\
&= \frac{2(nk + n^2k + n^2)}{4(2k+1)} - \frac{(2k+1)n^2}{4(2k+1)} \\
&= \frac{n(2k+n)}{4(2k+1)}.
\end{aligned}
$$

Therefore

$$
Var(X) = \lim_{\mu_. \to \infty} Var(Y_1 | Y_. = n; \mu_., k) = \frac{n(2k+n)}{4(2k+1)}.
$$

The Binomial variance analogue to $E(X) = n/2$ is $n/4$. The Conditional NB variance is inflated by the factor $(2k+n)/(2k+1)$. If k is small then the inflation is large, whereas if $k \to \infty$ the variance inflation goes to one.

Clearly the conditional probability is sensitive to the value of $\mu_.$. As such it is not directly useful for inferences about ρ. However, we explore two approximations for $\mu_.$ that we hope will lead to good inferences about ρ. We pursue the conditional approach because we anticipate that it will give more reliable estimates of k. Although k is a nuisance parameter, it will be important to have good estimates of k to get reliable confidence intervals for ρ.

The first approach we explore is to replace $\mu_.$ with n, which is the mle of $\mu_.$ as $k \to \infty$. The second approach we explore is to replace $\mu_.$ with its approximate mle based on the marginal distribution of $Y_. = n$, but treating p and k as fixed. This can be different from n when k is small. In effect we replace $\mu_.$ in (??) with a function of n, p and k. This is similar to the concentration approach we used for $\mu_.$ in Section 2.4.1.

The approximation we use for the mle of $\mu_.$ has a closed form expression. We require this because it makes the numerical calculation of the mle of p and k much more feasible. If we used numerical methods to compute $\mu_.$ then this would involve two levels of numerical optimization, optimize over $\mu_.$ for fixed p and k , and then optimize over p and k. The first numerical optimization for $\mu_.$ will introduce roughness into the likelihood surface for p and k which means that standard numerical derivative-based optimization routines (such as *optim* or *nlminb* in **R**) can be used to estimate p and k. Fortunately the mle for $\mu_.$ can be closely approximated using a Taylor's series expansion of $E(Y_1|Y_. = n; \mu_., k)$ about $\mu_. = n$. We present the expansion in the next section, followed by the approximate marginal mle of $\mu_.$.

2.4.4 Taylor Series Approximation of the conditional mean

The NB conditional expectation can be reasonably approximated using a Taylor's series expansion around $\mu_. = n$ (see shaded area in Figures -). We will use this result in the next section to develop a better conditional estimator of p. Let $\Psi(\mu_.) = E(Y_1|Y_. = n; \mu_., k)$ and let $\phi_x(\mu_.)$ denote the numerator term in the NB pdf,

$$\phi_x(\mu_.) = \frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)} \left(\frac{p}{q}\right)^x \left(\frac{q\mu_. + k}{p\mu_. + k}\right)^x . \tag{2.34}$$

By definition $\Psi(\mu_.) = \frac{\sum x\phi(\mu_.)}{\sum \phi(\mu_.)}$. It is not hard to show that

$$\begin{aligned}\frac{\partial\phi_x(\mu_.)}{\partial\mu_.} &= \left\{\frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)}\right\}\left\{\left(\frac{p}{q}\right)^x\right\}\left\{x\left(\frac{q\mu_. + k}{p\mu_. + k}\right)^{x-1}\right\} \\ &\times \left\{\frac{q(p\mu_. + k) - p(q\mu_. + k)}{(p\mu_. + k)^2}\right\}\end{aligned}$$

$$
\begin{aligned}
&= \left\{ \frac{\Gamma(k+x)\Gamma(k+n-x)}{\Gamma(x+1)\Gamma(n-x+1)} \right\} \left\{ \left(\frac{p}{q}\right)^x \right\} \left\{ x\frac{(q\mu_. + k)^{x-1}}{(p\mu_. + k)^{x+1}} \right\} \{k(q-p)\} \\
&= \frac{k(q-p)x\phi_x(\mu_.)}{(p\mu_. + k)(q\mu_. + k)}. \qquad (2.35)
\end{aligned}
$$

Let $\Psi_i(\mu_.) = \frac{\sum x^i \phi_x(\mu_.)}{\sum \phi_x(\mu_.)}$, i= 2,3. The first-order Taylor's series approximation of $\Psi(\mu_.)$ is

$$
\Psi(\mu_.) \doteq \Psi(n) + \left. \frac{\partial \Psi(\mu_.)}{\partial \mu_.} \right|_{\mu_. = n} (\mu_. - n), \qquad (2.36)
$$

where

$$
\begin{aligned}
\Psi'(\mu_.) &= \frac{\partial}{\partial \mu_.} \left\{ \frac{\sum x\phi_x(\mu_.)}{\sum \phi_x(\mu_.)} \right\} \\
&= \frac{\left\{\sum \phi_x(\mu_.)\right\}\left\{\sum x\phi_x'(\mu_.)\right\} - \left\{\sum x\phi_x(\mu_.)\right\}\left\{\sum \phi_x'(\mu_.)\right\}}{\left\{\sum \phi_x(\mu_.)\right\}^2} \\
&= \frac{\left\{\sum \phi_x(\mu_.)\right\}\left\{\sum x\frac{k(q-p)x\phi_x(\mu_.)}{(p\mu_.+k)(q\mu_.+k)}\right\} - \left\{\sum x\phi_x(\mu_.)\right\}\left\{\sum \frac{k(q-p)x\phi_x(\mu_.)}{(p\mu_.+k)(q\mu_.+k)}\right\}}{\left\{\sum \phi_x(\mu_.)\right\}^2} \\
&= \frac{\frac{k(q-p)}{(p\mu_.+k)(q\mu_.+k)}\left[\left\{\sum \phi_x(\mu_.)\right\}\left\{\sum x^2\phi_x(\mu_.)\right\} - \left\{\sum x\phi_x(\mu_.)\right\}^2\right]}{\left\{\sum \phi_x(\mu_.)\right\}^2} \\
&= \frac{\frac{k(q-p)}{(p\mu_.+k)(q\mu_.+k)}\left[\Psi_2(\mu_.)\left\{\sum \phi_x(\mu_.)\right\}^2 - \Psi^2(\mu_.)\left\{\sum \phi_x(\mu_.)\right\}^2\right]}{\left\{\sum \phi_x(\mu_.)\right\}^2} \\
&= \frac{k(q-p)}{(p\mu_. + k)(q\mu_. + k)} \left\{\Psi_2(\mu_.) - \Psi^2(\mu_.)\right\}.
\end{aligned}
$$

$\Longrightarrow$

$$
\Psi'(\mu_.) = \frac{\partial \Psi(\mu_.)}{\partial \mu_.} = \frac{k(q-p)}{(p\mu_. + k)(q\mu_. + k)}\{\Psi_2(\mu_.) - \Psi^2(\mu_.)\}.
$$

Hence, the first-order approximation is

$$
\Psi(\mu_.) \doteq \Psi(n) + \frac{k(q-p)(\mu_. - n)}{(pn + k)(qn + k)}\{\Psi_2(n) - \Psi^2(n)\}. \qquad (2.37)
$$

The first-order approximations are shown as the dashed straight lines in Figures [illegible]

To derive the second-order approximation, we first evaluate

$$
\begin{aligned}
\Psi''(\mu_.) = \frac{\partial^2\Psi(\mu_.)}{\partial\mu_.^2} &= \frac{k(q-p)}{(p\mu_. + k)(q\mu_. + k)}\{\Psi_2'(\mu_.) - 2\Psi_2(\mu_.)\Psi_2'(\mu_.)\} \\
&+ \frac{-k(q-p)p(q\mu_. + k)q(p\mu_. + k)}{(p\mu_. + k)^2(q\mu_. + k)^2}\{\Psi_2(\mu_.) - \Psi^2(\mu_.)\} \\
&= \frac{-k(q-p)((pq\mu_.)^2 + 2pqk\mu_. + k^2)}{(p\mu_. + k)^2(q\mu_. + k)^2}\{\Psi_2(\mu_.) - \Psi^2(\mu_.)\} \\
&+ \frac{k(q-p)}{(p\mu_. + k)(q\mu_. + k)}\{\Psi_2'(\mu_.) - 2\Psi(\mu_.)\Psi'(\mu_.)\},
\end{aligned}
$$

where

$$
\Psi_2(\mu_.) = \frac{\sum x^2\phi_x(\mu_.)}{\sum \phi_x(\mu_.)}
$$

and

$$
\begin{aligned}
\frac{\partial\Psi_2(\mu_.)}{\partial\mu_.} &= \frac{\{\sum \phi_x(\mu_.)\}\{\sum x^2\phi_x'(\mu_.)\} - \{\sum \phi_x'(\mu_.)\}\{\sum x^2\phi_x(\mu_.)\}}{\{\sum \phi_x(\mu_.)\}^2} \\
&= \frac{\{\sum \phi_x(\mu_.)\}\{\sum \frac{k(q-p)}{(p\mu_.+k)(q\mu_.+k)}x^3\phi_x(\mu_.)\} - \{\frac{k(q-p)\sum x\phi_x(\mu_.)}{(p\mu_.+k)(q\mu_.+k)}\}\{\sum x^2\phi_x(\mu_.)\}}{\{\sum \phi_x(\mu_.)\}^2} \\
&= \frac{\frac{k(q-p)}{(p\mu_.+k)(q\mu_.+k)}[\Psi_3(\mu_.)\{\sum \phi_x(\mu_.)\}^2 - \Psi(\mu_.)\Psi_2(\mu_.)\{\sum \phi_x(\mu_.)\}^2]}{\{\sum \phi_x(\mu_.)\}^2},
\end{aligned}
$$

which implies that

$$
\Psi_2'(\mu_.) = \frac{\partial\Psi_2(\mu_.)}{\partial\mu_.} = \frac{k(q-p)}{(p\mu_. + k)(q\mu_. + k)}\{\Psi_3(\mu_.) - \Psi_2(\mu_.)\Psi(\mu_.)\}.
$$

These results are used to compute the second-order approximation,

$$
\Psi(\mu_.) \doteq \Psi(n) + (\mu_. - n)\Psi'(n) + \frac{1}{2}(\mu_. - n)^2\Psi''(n), \tag{2.38}
$$

which are shown as dotted-dashed curves in Figures [illegible] - [illegible] .

Within a neighborhood of ±50% of $\mu_. = n$ the first and second order approximations work well. We will use these in the next section to find a closed-form expression for the approximate mle of $\mu_.$, given p and k, based on $Y_. = n$.

2.4.5 Estimation of $\mu_.$

An approach to deal with the $\mu_.$ term in (??) for inferences about ρ is to replace $\mu_.$ by its marginal mle but with ρ and k treated as unknown parameters. The marginal pdf for $Y_.$ is given by

$$\begin{aligned}
P(Y_. = n) &= \sum_{x=0}^{n} P(Y_1 = x)P(Y_2 = n - x) \\
&= \sum_{x=0}^{n} \frac{\Gamma(k+x)}{\Gamma(k)\Gamma(x+1)} \left(\frac{\mu_1}{\mu_1 + k}\right)^x \left(\frac{k}{\mu_1 + k}\right)^k \frac{\Gamma(k+n-x)}{\Gamma(k)\Gamma(n-x+1)} \left(\frac{\mu_2}{\mu_2 + k}\right)^{n-x} \\
&\times \left(\frac{k}{\mu_2 + k}\right)^k \\
&= \frac{k^{2k}}{\Gamma(k)^2} \left(\frac{q\mu_.}{q\mu_. + k}\right)^n (p\mu_. + k)^{-k}(q\mu_. + k)^{-k} \sum_{x=0}^{n} \phi_x(\mu_.),
\end{aligned}$$

where $\phi_x(\mu_.)$ is given by ([illegible]). The loglikelihood for $\mu_.$, $\Lambda(\mu_.)$, is given by

$$\Lambda(\mu_.) = C + n\log(\mu_.) - k\log(p\mu_. + k) - k\log(q\mu_. + k) + \log\left\{\sum_{x=0}^{n} \phi_x(\mu_.)\right\},$$

where C denotes terms that do not involve $\mu_.$. After some simplification,

$$\frac{\partial\Lambda(\mu_.)}{\partial\mu_.} = n(p\mu_. + k) - (2pq\mu_. + k)\mu_. + (q - p)\mu_.\Psi(\mu_.). \tag{2.39}$$

To derive this expression we used ([illegible]) and the fact that

$$\frac{\partial}{\partial\mu_.}\log\left\{\sum_{x=0}^{n}\phi_x(\mu_.)\right\} = \frac{\sum_{x=0}^{n}\phi_x'(\mu_.)}{\sum_{x=0}^{n}\phi_x(\mu_.)} = \frac{k(q-p)\Psi(\mu_.)}{(p\mu_. + k)(q\mu_. + k)}.$$

The first-order approximate mle of $\mu_.$ can be obtained by replacing $\Psi(\mu_.)$ in ([illegible]) with (2.36). This yields a quadratic equation in $\mu_.$, $a_1\mu_.^2 + a_2\mu_. + a_3$, where

$$\begin{aligned}
a_1 &= (q-p)\Psi'(n) - 2pq, \\
a_2 &= np + (q-p)\{\Psi(n) - n\Psi'(n)\} - k, \\
a_3 &= nk.
\end{aligned}$$

The root of this equation is easy to obtain. It can be shown that the first-order approximate mle for $\mu_.$ is n when $p = 1/2$.

A slightly more accurate approximation for the mle of $\mu_.$ may be obtained by replacing $\Psi(\mu_.)$ in ([illegible]) with ([illegible]). However, in simulations we did not find any advantage in using the second-order approximation for estimating ρ so we do not present further details about this approximation. However, it is more complicated because the approximate mle is the root of a cubic polynomial.

We illustrate the accuracy of the approximations to the mle in Figures [illegible] - [illegible]. The profile values are $2\times$ the difference between the maximum log-likelihood and the log-likelihood for values of $\mu_.$; hence, the minimum of the profile is zero. The second-order approximate mle for $\mu_.$ is usually coincident with the mle, except when k=0.5 and p=0.1. The first-order mle is usually very close to the mle as well. Note that the mle can be quite different from n when k and p are small. The likelihood profiles are flat indicating that a wide range of $\mu_.$ values are consistent with a specific value of n. Based on a χ^2_1 distribution, a 95% confidence interval for $\mu_.$ would include values such that the profile was less than $\chi^2_{1,0.95}$=3.84. Such confidence intervals would greatly exceed the range of $\mu_.$'s in most panels of Figures [illegible] - [illegible]. This suggests that, as expected, a single observation of $Y_.$ does not give much information about $\mu_.$.

The percent differences in the first-order mle ($\tilde{\mu}_.$),

$$\% difference = 100\left(\frac{\tilde{\mu}_. - n}{n}\right), \tag{2.40}$$

are shown in Figure [illegible]. When k and p are small then the mle can be almost 50% greater than n; however, if $k \geq 1$ and $p \geq 0.25$ then the differences are less than 8%. This figure suggests that the mle converges to n as $k \to \infty$ or $p \to 0.5$.

2.4.6 Models Diagnostics

Once a model has been fitted to the observed values of response variables, it is essential to check that the fitted model is actually valid.

There are a number of ways in which a fitted model may be inadequate. The most serious of these is that the linear systematic component of the model may be incorrectly specified; for example, it may not include explanatory variables that really should be in the model. Also, the data may contain particular observations, termed outliers, that are not well fitted by the model. Finally, the assumption that the observed data come from a particular probability distribution, for example, the Negative Binomial distribution, may not be valid.

The techniques used to examine the adequacy of a fitted model are known as diagnostics. These techniques frequently involve statistics that are based on differences between the fitted values under a model and the observations to which that model has been fitted.

Measures of agreement between an observation of a response variable and the corresponding fitted value are known as residuals. These quantities can provide much information about the adequacy of a fitted model.

In this thesis, we use a common form of residuals known as Pearson chi-square residuals. The chi-square residual are defined by

$$r_i = \frac{y_i - E(Y_i)}{\sqrt{Var(Y_i)}}. \tag{2.41}$$

It is simply the residuals scaled by the estimated standard deviation of Y.

Suppose that a linear logistic model is fitted to n observations of the form y_i/n_i, $i = 1, 2, ..., N$, and the corresponding fitted value of y_i is $\hat{y}_i = n_i\hat{p}_i$. Then for the ith observation, the Binomial chi-square residual is given by

$$r_i = \frac{y_i - n_i\hat{p}_i}{\sqrt{n_i\hat{p}_i\,(1-\hat{p}_i)}}, \tag{2.42}$$

and the over-dispersed binomial chi-square residual is

$$r_i = \frac{y_i - n_i \hat{p}_i}{\sqrt{\phi n_i \hat{p}_i (1 - \hat{p}_i)}}. \tag{2.43}$$

Equation ( ) is also used in this thesis to compute residuals for both the Conditional NB and the Concentrated NB models. For the Conditional NB model, the conditional expectation and conditional variance need to be found based on ( ) and ( ) so that equation ( ) is expressed as

$$r_i = \frac{y_{i1} - \sum_{y_1=0}^{n_i} y_1 p_{n_i}(y_1)}{\sqrt{\sum_{y_1=0}^{n_i} y_1^2 p_{n_i}(y_1) - \left\{ \sum_{y_1=0}^{n_i} y_1 p_{n_i}(y_1) \right\}^2}}, \tag{2.44}$$

where $p_{n_i}(y_1) = P(Y_1 = y_1 | Y_. = n_i)$ as given in (??)

In Concentrated NB model, residuals are computed for both test and control observations. The NB chi-square residuals are

$$r_{ij} = \frac{y_{ij} - \hat{\mu}_{ij}}{\sqrt{\hat{\mu}_{ij} \left(1 + \hat{\mu}_{ij}^2 / \hat{k}\right)}}, \quad i = 1, ..., N, \; j = 1, 2. \tag{2.45}$$

2.4.7 Computer Software for Models Analysis

R provides a powerful interactive computing environment for data analysis with extensive graphical facilities. This package is not straightforward to use as other packages. The scope of the package can be extended by writing new functions or modifying existing ones. The software incorporates a function named *glm* used to fit generalized linear models, and for binary data analysis the argument *family=binomial* is included. Having fitted a model using the function *glm*, the function *summary* is used to obtain parameter estimates and their standard errors. The function *residuals* can be used to calculate the values of certain residuals. Numerical solutions for maximizing likelihood functions can be found by *nlminb* and *optim* functions.

Chapter 3

Real Data Analysis

3.1 Comparative Fishing Survey Protocols

The main objective of the comparative fishing exercise was to determine if differences exist between WT and AN catchabilities (q) when both vessels used the standard survey trawl. Data from paired tows were collected to quantify potential differences. The location of the comparative fishing was off the east coast of Newfoundland, in Northwest Atlantic Fisheries Organization (NAFO) Subdivision 3Ps, and Divisions 3LNO (see Figure 1.[illegible]). Tow stations were selected randomly as part of research surveys. High density aggregations were not specifically targeted because information was required on differences in catchability when stock densities are high and low.

The WT followed normal survey protocols, and the AN surveyed for comparison purposes only. The vessels were instructed to tow on the same course, and the WT relayed the course to the AN. On slope edges, where side by side tows were not feasible due to depth differences, one vessel towed ahead of the other, alternating the lead vessel on a tow-by-tow basis. This was done so that the end of the tow for the trailing vessel occurred at a position just before the start of the tow for the leading vessel; that is, there was no overlap in the area covered by the tows. The same depth range for each paired tow was maintained as close as possible between boats.

Differences in the depths fished for both vessels were minimized, for a target of less than 10% during comparative tows. If the WT had an unsuccessful set, both vessels repeated their tows, moving slightly so that the same grounds were not towed over again. A full range of Scanmar trawl geometry sensors were used on each vessel, including n additional depth sensor mounted 50 m in front of the trawl doors. This additional sensor monitored the effect of using trawl warp on the AN which was 1/8 inch larger in diameter than that used on the WT (1 inch).

Let y_{is} be the number of fish caught at the ith tow station by vessel s. We refer to the replacement vessel as s = t for the test vessel, and we refer to the vessel to be replaced as s = c for control. We assume that the replacement vessel is the AN, although our results can easily be adjusted if the WT is the replacement vessel.

3.2 Results

A total of 57 paired survey sets with the AN and WT were carried out in 2005. There were 49 sets with no catch for both vessels, so these sets provide no information about rho. The number of sets with some catch by either vessel was 106 - 49 = 57. Most major commercial species had some survey coverage. Sets were located in the far offshore portion of 3Ps (Fig.), the shelf area in 3N, and the northern part of 3L. Note that in this figure the location of the plotting symbols indicates the average location of fishing for the two vessels. The size, type, and color of the plotting symbols give information about the within-pair differences in catches (see figure caption). The actual catches are shown in Table C.1. The distance between paired tows was relatively constant, with a maximum of 3.7 km. Tow depths were also usually similar, with a maximum absolute difference of 37 m (Cadigan et al, 2006). Most of the Witch flounder catches occurred in Subdivision 3Ps. In Fig. we show the difference in catches for each pair of tows, scaled by their "Poisson" standard deviation, which was the square root of the sum of the catches. Black symbols merely indicate potential outliers; however, these "residuals" were not adjusted for over-dispersion, and we do not suggest that these catches are outliers.

In this thesis, a length effect is not considered. That is, relative efficiency is assumed to be the same for all lengths, and only affected by vessels (WT and AN). This implies that equations (.) and ( ) can be simplified to

$$\frac{e^{\beta}}{1+e^{\beta}} = p, \tag{3.1}$$

and

$$logit(p) = \beta + u. \tag{3.2}$$

Table C.2 shows a summary of results for all fitted models. Appendix B.1 shows the results of fitting the Binomial logistic models. Substantially more user-developed computer code is required to fit the NB conc. and NB cond. models. This code is not given in Appendix B.

3.2.1 Binomial Model

The standard binomial logistic regression model was implemented in **R** using the function *glm*. Appendix B.1 gives the code results for this model. The estimate of β is -0.17214. That is, relative efficiency estimate is exp(-0.17214) = 0.841. Note that the estimate sign is negative indicating that WT had a slightly lower relative efficiency than the AN. The 95% confidence interval does not cover one, which leads to the conclusion that the vessels had significantly different catchabilities (q's). The data and estimated relative efficiency are shown in Figure (solid line) as a straight-line through the origin with slope $\hat{\rho}$ where $\hat{\rho} = exp(\hat{\beta})$. Residuals of this model are presented in Figure . More than 5% of the chi-square residuals have absolute value > 2, which suggests over-dispersion.

3.2.2 Over-dispersed Binomial model

Appendix B.2 gives the code results for the Overdispersed Binomial model. It was implemented in **R** in which the quasi-likelihood approach was used to estimate the over-dispersion parameter. Note that the parameter estimates from this approach are identical to those in the previous section (see Table C.2), but the standard error is larger. This leads to a wider confidence interval for the vessel effect. The 95% confidence interval for ρ also does not cover one, which leads to the same conclusion from the Binomial model with no over-dispersion. The relative efficiency from the Overdispersed Binomial model is presented in [illegible]. The chi-square residuals adjusted for over-dispersion are shown in Figure [illegible]. Note that their overall magnitude is smaller than the Binomial residuals (Fig. [illegible]). Figure [illegible] shows that one of the sets had a residual of value less than -4. This set may be an outlier.

3.2.3 Mixed Binomial model

R code and results for the Mixed Binomial model are presented in Appendix B.3. The maximization was implemented in **R** using the function *glmmML*. The estimate of relative efficiency (Table C.2) from this model differs somewhat from what we obtained with the previous two models. In this model, $\hat{\beta}$ = -0.1251, which implies that $\hat{\rho}$ = exp(-0.1251) = 0.884. The 95% confidence interval for ρ covers one, which leads to the conclusion that the vessels did not have significantly different catchabilities. This is different than the conclusion from the Binomial and over-dispersed Binomial models. Figure [illegible] shows the relative efficiency. Residuals for this model are not presented in this thesis because the R *glmmML* procedure does not produce any residuals. Chapter 5 gives more discussion about residuals for the GLMM model.

3.2.4 Conditional Negative Binomial, $\mu_. = n$

The conditional NB model with $\mu_. = n$ (total observed catch) was estimated by maximum likelihood using the **R** function *optim*. The resulting β estimate was similar to the Mixed Binomial model estimate. Relative efficiency for this model is presented in Figure [illegible]. Residuals are presented in Figure [illegible]. All residual values were within -3 and 3, and there is no evidence of model mis-specification.

3.2.5 Conditional Negative Binomial, $\mu_.$ estimated

The estimate of β, from the conditional NB model in which $\mu_.$ was replaced by its marginal estimate, was also similar to the Mixed Binomial model estimate. Relative efficiency for this model is presented in Figure [illegible]. Residuals are presented in Figure [illegible]. All residual values were within -2 and 2, and there is no evidence of model mis-specification.

3.2.6 Full Negative Binomial

The concentrated NB model likelihood was maximized using the **R** function *optim*. It produced a β estimate that is very close to those we obtained from GLMM, Conditional NB ($\mu_. = n$) and Conditional NB ($\mu_.$ estimated). Also, it produced a standard error that is clearly smaller than standard errors produced by GLMM and Conditional NB (for both $\mu_.$ cases) models and this resulted in having a shorter 95% confidence interval. Hence, the conclusion from this model is that relative efficiency was significantly different from one, whereas from the GLMM and conditional NB model analyses we did not conclude that $\rho \neq 1$. Estimated relative efficiency is shown in Figure [illegible]. Also, residuals of this model are presented for both of test and control catches in Figure [illegible]. The residuals look reasonable.

3.3 Conclusions

The results in Table C.2 suggest that there were no significant difference in catchability for the WT and AN fishing the Witch Flounder trawl based on the GLMM, Conditional NB ($\mu_.$ estimated) and Conditional NB ($\mu_. = n$) models. This is because β confidence intervals from those models contained zero as a value. That is, in the GLMM, Cond NB and Conc NB models, β was not significantly different from zero. The analyses for the Binomial, over-dispersed Binomial, and Full NB models suggested that relative efficiency was significantly different from one. However, in the next chapter we show that confidence intervals from the latter three approaches are much less reliable than the GLMM and Conditional NB models. Hence, our conclusion is that the vessels did not have significantly different catchabilities for Witch flounder.

Chapter 4

Simulation Study

4.1 Design

Simulated data were generated to compare estimates of β from the OD Bin, GLMM, Conc NB, and Cond NB models. Data were generated using pseudo-random number generators for the Negative Binomial distribution. In each data set, paired catches were generated from independent Negative Binomial distributions with parameters β, $\mu_.$, and k. Recall that the NB means for each pair, μ_1 and μ_2, are obtained as $mu_1 = p\mu.$ and $mu_2 = q\mu.$, where $q = 1 - p$ and $logit(p) = \beta$. The simulation parameter values were: $\beta = 0, 0.4, 0.69, 1.60$ (corresponding ρ values described below), $\mu. = 15, 30, 60$ and $k = 1, 3, 10$. In the simulation sets of net catch sample data were randomly generated as outlined above. Sample sizes (i.e. number of sets) of $N = 20$, 35 and 50 were considered. The simulation values of β, $\mu.$, k, and n were chosen to cover the range of values that might occur in real comparative fishing data sets. Two thousand pairs of data sets were generated for each of the 108 possible combinations of β value, total mean value $\mu.$, k and sample size. The corresponding ρ true simulation values are 1, 1.5, 2 and 5.

Bias, standardized bias, mean square error (MSE) and confidence intervals coverage were assessed. The total bias of the estimated parameter $\hat{\beta}$ is given by

$$Total\ Bias = \frac{1}{2000}\sum_{i=1}^{2000}(\hat{\beta}_i - \beta), \tag{4.1}$$

where 2000 is the number of iterations performed for any one of the previous mentioned simulation schemes. The total standardized bias is given by

$$Total\ Standardized\ Bias = \frac{1}{2000}\sum_{i=1}^{2000}\frac{(\hat{\beta}_i - \beta)}{SE(\hat{\beta}_i)} \times 100\%. \tag{4.2}$$

Also, ( ) can used for computing total bias for the estimated overdispersion parameter ($\hat{k}$) in which β is replaced by k.

Confidence interval coverage is also important to understand for reliable statistical inferences. We evaluated the accuracy of 90% and 95% confidence intervals computed as follows

$$\hat{\beta} \pm Z_{\alpha/2}SE(\hat{\beta}). \tag{4.3}$$

Simulated confidence interval coverage is the proportion of simulations (i.e. out of 2000 for each set of simulation parameters) in which the true simulation value of β falls within the computed confidence interval endpoints. For reliable confidence intervals the proportion of (1-α)% coverage confidence intervals covering the true β value should be close to (1-α)%.

The conditional NB model was estimated using three different options for $\mu_{\cdot}$, namely 1) $\mu_{\cdot}$ estimated; 2) $\mu_{\cdot} = n$; and 3) $\mu_{\cdot}$ fixed at the true simulation value. The latter approach could not be used in practise because we would not know the true value for $\mu_{\cdot}$; however, it is useful for understanding problems with the cond NB approach associated with not knowing $\mu_{\cdot}$.

4.2 Simulation Outcomes

Simulation results are shown in Table C.3 C.29. Columns 4 – 9 are for the following 6 estimators: 4) Conditional NB with mu.dot = est. 5) Conditional NB with mu.dot = n, 6) Conditional NB with mu.dot = true, 7) Concentrated NB, 8) GLIM, 9) GLMM. Results for all simulation schemes are better summarized in figures [illegible] to [illegible]. Biases are presented in Figure [illegible]. These results are also shown in Table C.3, C.4 and C.5. Mean Square Error (MSE) is presented in Figure [illegible] and Tables C.9, C.10 and C.11. Simulated lower, upper and total coverage of 90% and 95% of β estimates can be seen in Figures [illegible] to [illegible] and Tables C.12 to C.29. In each figure, nine panels are given so that each panel shows the results for a combination of n and k for the previous mentioned 6 estimators, which are shown at the top and right-hand side, respectively. The model and value for μ. are shown at the left-hand side. Each group of points corresponds to a value of $\rho = exp(\beta) = 1$ (top line), 1.5, 2, and 5 (bottom line). Conditional NB with mu.dot = est is abbreviated as Cond1, Conditional NB with mu.dot = n as Cond2 and Conditional NB with mu.dot = true as Cond3. Also, Concentrated NB is abbreviated as Conc. Zero is shown as a solid vertical line.

The bias was generally found to fall within -0.05 and 0.05 values except for β = 1.6 ($\rho = 5$) in the Conditional NB model where some extreme negative bias values were found. Those negative low bias values can be clearly seen when k = 1 or 3. As the overdispersion parameter (k) gets larger then the six estimators give smaller bias values. MSE results are presented in Figure [illegible]. MSE decreases as k increases (i.e. over-dispersion decreases) for all models. Also, it is clear that when k and ρ increase then the MSE's are smaller, for all estimators under study. The combination of bias and MSE values can be found in Figures [illegible] to [illegible] where the total 90% and 95% coverage are presented. In Figure [illegible], it can be clearly seen that when k = 1 both the Concentrated NB and GLIM estimators produce coverage values less than 80% and Cond1, Cond2, Cond3 and GLMM produce coverage values closer to 90%.

If $k = 3$ or 10, then the Concentrated NB estimator produces coverage values less than 80% while the GLIM coverages tend to get higher than than 80% and closer to 90%. In Figure [illegible], the Concentrated NB estimator resulted in the lowest coverage values. The 95% coverage values are generally less than 95% for all k combinations. Note that Concentrated NB estimator produced coverage values around 85% for all k combinations. All Cond1, Cond2, Cond3 and GLMM give coverage values were close to 95% for all k combinations while GLIM coverage values increased when k got larger.

The general conclusion from the simulations is that bias and standardized bias in the various estimators of β tended to be small, except when $\rho = 5$. In this case the conditional NB approaches with $\mu.$ estimated or $\mu. = n$ had substantial bias, which was worse when $\mu. = n$. This bias did not occur when $\mu.$ was fixed at the true value which suggests that the problem when $\rho = 5$ is related to the unknown $\mu.$ parameter in the conditional likelihood. These biases did not affect MSE. In fact, in all three conditional NB models the MSE was usually lowest when $\rho = 5$. Total coverage of confidence intervals was affected by the bias, especially when $\mu. = n$. Total coverage was reasonably accurate when $\mu.$ was estimated, although one-tailed coverage tended to be less accurate when $\rho = 5$. All estimators yielded confidence interval coverage that were somewhat lower than the nominal considered percent level. The Concentrated NB estimator performed very poorly in all cases. The Binomial approach resulted in poor confidence coverage, especially for small $k = 1, 3$.

The GLMM produced confidence intervals that had relatively good coverage properties, especially when $n = 50$. However, this procedure was clearly inefficient in terms of MSE when $k = 1$, which is a practically relevant amount of over-dispersion.

Chapter 5

Conclusion

The results in Chapter 3 suggested that there were small differences in catchability between the WT and AN for the species Witch flounder. The sign of the log relative efficiency parameter estimate for different models under study was always negative which provides some additional evidence that the catchability of the AN was lower than the WT. However, the effect, if it exists, appeared small and could be ignored without serious consequences.

The Conditional Poisson model (i.e. Binomial model) suggested that the AN vessel is 84% as efficient as the WT vessel, and the effect was statistical significant. However, over-dispersion was apparent in this model. An analysis of Pearson chi-square residuals showed that many were $\geq \pm 2$, more than one would expect due to simply random variability. The data exhibited more variability than could be explained by Binomial sampling.

One source of over-dispersion in the Binomial model is the erroneous assumption that differences in stock densities fished by each trawler were identical. In practise this does not happen, although differences in stock densities at each tow site within a pair should be completely random so that they can be viewed as iid samples from some distribution of densities. We conjecture that this was the motivation by Benoit and Swain (2003) and Lewy et al. (2004) for using a Binomial over-dispersed parameter. We also fitted a Binomial model with an over-dispersed variance, using quasi-likelihood to estimate the over-dispersion parameter. The estimate of β was identical to the Binomial model estimate, but the standard error was larger and confidence intervals were wider. We suggest the wider confidence intervals are more accurate. Nonetheless, the results still suggested that β was marginally significantly different from zero. The residuals from the over-dispersed Binomial looked more reasonable as well. Benoit and Swain (2003) suggested that the over-dispersed Binomial approach still led to false significance, and they used a randomization method for determining statistical significance. The main research in this thesis was to investigate more thoroughly models for a specific type of over-dispersion, which is within-pair random differences in stock densities.

Three alternative methods were explored. First, we analyzed the data using a mixed binomial model with an independent and identically distributed random standard normal effect for each set. The mixed binomial model (GLMM) suggested that the AN was 88% as efficient as WT vessel, but that ρ was not significantly different from one at the 5% level. A well developed theory exists for residual diagnostics in fixed effects models, but much less seems to be available for mixed-effects models and we could not produce residuals for the GLMM. Residual diagnostics for GLMM's are practically important and future research in this area would be useful. Simulation results showed that this model did a good job in avoiding bias and giving confidence interval coverage close to the nominal desired levels; however, it performed fairly poorly in terms of MSE for the type of over-dispersion we considered. This model seems more reliable for practical use than the over-dispersed Binomial model approach; however, further improvements are possible.

We also investigated the problem of overdispersion by assuming that the vessel's catches have a Negative Binomial distribution, which can be regarded as a generalization of the Poisson distribution with an additional parameter allowing the variance to exceed the mean. Direct estimation of the NB model by maximum likelihood suggested that the AN vessel was 88% as efficient as the WT vessel, and that the vessel effect was marginally significant. However, in our simulation analysis we found that this model was the worst among those examined, in terms of confidence interval coverage The problem is related to gross over-estimation of the NB k parameter, which leads to gross under-estimation of standard errors and confidence intervals that are much too narrow. Direct estimation of variance parameters by maximum likelihood is known to be biased when there are many mean parameters, and this is certainly the case for paired-trawl calibration data. In this model, the many $\mu_i'.s$ are nuisance parameters and some adjustment for estimating these parameters is required for inferences about the NB k parameter. Our simulation results showed that the mle of k was badly biased. Using the full NB model is not recommended. This is why we investigated conditioning for inferences about relative efficiency.

Conditioning with paired count data when the data are NB distributed has received little study in the statistical literature, and this was an important contribution of this thesis. The conditional approach is commonly used when data are Poisson distributed, and a good way to deal with the pair-total nuisance parameters, $\mu.$'s. However, the approach is more problematic with NB data because the $\mu.$'s are not eliminated in the conditional distribution. We explored two options for dealing with these nuisance parameters. One was to replace them by n, which is their mle when the NB over-dispersion is small (i.e. k is large), and the other option was to replace them with their direct mle based on the marginal distribution of the paired-totals.

The conditional NB model performed slightly better in our simulations than the Mixed Binomial model in terms of MSE; however, the conditional approaches we investigated had bias problems when the vessel effect was large, and this led to poorer confidence interval coverage. The option of estimating $\mu.$ gave better bias results and good total confidence interval coverage, but there was still some problems with one-sided coverages.

The conditional NB model suggested that AN vessel is 88% as efficient as WT. This model produced a standard error that was appreciably larger than those for the models under except GLMM model. This is because it allows for overdispersion. Residuals produced in Conditional NB are presented in Figure when $\mu_{.} = n$ and in Figure when $\mu_{.}$ is estimated. It is clearly noted that if $\mu_{.} = n$ then the obtained residuals are relatively smaller than those obtained when $\mu_{.}$ is estimated in the Conditional NB model. In general, residuals obtained under both cases of Conditional NB model are relatively small compared to those produced with other models. In this model, residuals values records are within -3 to +3, while it can be noted in figures and that some residuals values fell outside the range -3 to +3. For both options of dealing with $\mu_{.}$, ρ was not significantly different from one at the 5% level because confidence intervals for β estimates contained zero as a value.

In this thesis we demonstrated that the Conditional NB and Mixed Binomial models performed better than the Binomial logistic model with over-dispersion or the full NB model. However, the efficacy of the these approaches for estimating relative efficiency requires further research especially for large β values. Also, we conclude that the full NB model and over-dispersed Binomial model provide poor confidence intervals and are not recommended for paired-trawl calibration studies in which within-pair gamma-type variations in local stock densities occur, which would seem to be common.

Appendix A

Evaluation of Conditional NB Probability

A.1

To compute the pdf, let $p_n(x) = P(Y_1 = x | Y_. = n)$

$$r_n(x) = \frac{p_n(x)}{p_n(x-1)} = \left(\frac{p}{q}\right)\left(\frac{q\mu_. + k}{p\mu_. + k}\right)\left(\frac{x+k-1}{x}\right)\left(\frac{n-x+1}{n-x+k}\right)$$

Let $c_n(x)$ be a numerator term in $p_n(x)$:

$$\begin{aligned}
c_n(0) =& \frac{\Gamma(k)\Gamma(k+n)}{\Gamma(n+1)}. \\
c_n(1) =& c_n(0)r_n(1) \\
c_n(2) =& c_n(1)r_n(2) = c_n(0)r_n(1)r_n(2) \\
c_n(3) =& c_n(2)r_n(3) = c_n(0)r_n(1)r_n(2)r_n(3) \\
& \quad etc.
\end{aligned}$$

This is an easy recursive formula to compute

$$p_n(x) = \frac{c_n(x)}{\sum_{x=0}^{n} c_n(x)}.$$

An even better formula is

$$\begin{aligned}
p_n(x) =& \frac{c_n^*(x)}{\sum_{x=0}^{n} c_n^*(x)}, c_n^*(x) = \frac{c_n(x)}{c_n(0)}, \\
c_n^*(1) =& 1, \\
c_n^*(1) =& r_n(1) \\
c_n^*(2) =& r_n(2) = r_n(1) r_n(2) \\
c_n^*(3) =& r_n(3) = r_n(1) r_n(2) r_n(3) \\
& etc.
\end{aligned}$$

Appendix B

R Software Codes

B.1

```
 Binfit <- glm(cbind(test,control) ~ 1, family=binomial,data=catches)
 summary(Binfit)
Call: glm(formula = cbind(test, control) ~ 1, family = binomial,
data = catches) Deviance Residuals:
    Min        1Q    Median       3Q       Max
-9.4286  -1.0936   0.2656   1.2513   7.6943 Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.17214    0.02923   -5.89 3.87e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
    Null deviance: 334.55  on 56  degrees of freedom
Residual deviance: 334.55  on 56  degrees of freedom AIC: 537.49
Number of Fisher Scoring iterations: 3
```

B.2

```
 Binfit.od <- glm(cbind(test,control) ~ 1,
family=quasibinomial(link = "logit"),data=catches)
 summary(Binfit.od)
Call: glm(formula = cbind(test, control) ~ 1, family =
quasibinomial(link = "logit"),
    data = catches)
Deviance Residuals:
    Min        1Q    Median       3Q       Max
-9.4286  -1.0936   0.2656   1.2513   7.6943 Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.17214    0.06891  -2.498   0.0154 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for quasibinomial family taken to be 5.557919)
    Null deviance: 334.55  on 56  degrees of freedom
Residual deviance: 334.55  on 56  degrees of freedom AIC: NA

Number of Fisher Scoring iterations: 3
```

B.3

```
 set<-catches$total
 Binfit.ri <- glmmML(cbind(test,control) ~ 1, family=binomial,data=catches,
 cluster = set)
Warning message: non-integer #successes in a binomial glm! in:
eval(expr, envir, enclos)
 summary(Binfit.ri)
```

```
Call:  glmmML(formula = cbind(test, control) ~ 1, family = binomial,
data = catches, cluster = set)
               coef se(coef)      z Pr(>|z|)
(Intercept) -0.1251   0.1218 -1.028    0.304

Standard deviation in mixing distribution:  0.7049 Std. Error:
0.09519

Residual deviance: 173.8 on 55 degrees of freedom AIC:177.8
```

Appendix C

Tables

Table C.1: Catch Summaries for Witch Flounder

$TestCatch$	$ControlCatch$	$TestCatch$	$ControlCatch$
1	0	113	125
0	1	141	119
0	1	64	64
1	0	95	79
13	18	146	145
1	0	92	48
0	1	58	83
1	0	56	5
1	0	27	19
2	0	13	13
5	3	68	73
1	1	1	5
18	33	0	1
93	310	1	3
55	23	10	14
17	23	41	55
95	105	89	82
4	13	42	36
1	3		
13	17		
4	11		
12	18		
102	112		
5	42		
5	7		
6	1		
18	30		
71	44		
74	119		
72	90		
46	67		
5	2		
59	131		
1	0		
65	84		
65	71		
41	46		
51	33		
75	132		

Table C.2: Numerical Results in Real Data Application. β is the logarithm of relative efficiency

$Model$	$\hat{\beta}$	$\hat{\rho}$	$Std.Error$	$\hat{\beta}$ 95%CI	$\hat{\rho}$ 95%CI
$CondPoi$	-0.17214	0.841	0.02923	$(-0.229, -0.114)$	$(0.795, 0.894)$
$ODBin$	-0.17214	0.841	0.06891	$(-0.307, -0.037)$	$(0.735, 0.963)$
$GLMM$	-0.12218	0.884	0.09290	$(-0.304, \quad 0.059)$	$(0.737, 1.060)$
$ConcNB$	-0.12463	0.882	0.06189	$(-0.245, -0.003)$	$(0.782, 0.997)$
$CondNB$ $(\mu_. = est.)$	-0.12604	0.881	0.09320	$(-0.308, \quad 0.056)$	$(0.734, 1.057)$
$CondNB$ $(\mu_. = n)$	-0.12604	0.881	0.09316	$(-0.308, \quad 0.056)$	$(0.734, 1.058)$

Table C.3: Bias (×100) in estimates of β from the simulations with $k = 1$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	0.27	0.85	1.18	0.22	−0.38	0.33
1.5	15	20	1.01	0.42	−0.60	−1.62	−1.54	1.63
2	15	20	1.41	0.10	1.02	−2.13	−0.35	3.01
5	15	20	−4.23	−16.06	0.83	−3.63	−1.06	4.68
1	30	20	−0.73	0.28	0.79	−0.65	−0.37	−0.69
1.5	30	20	2.00	0.04	1.48	0.67	1.47	2.60
2	30	20	0.42	−3.63	0.87	−0.82	−0.41	1.82
5	30	20	−8.29	−18.25	1.24	−3.58	−1.20	1.06
1	60	20	−0.85	−0.78	−0.95	−0.87	0.35	−0.69
1.5	60	20	−0.41	−1.57	2.74	−0.84	−0.37	0.03
2	60	20	0.54	−4.07	2.28	0.55	1.11	2.05
5	60	20	−7.55	−21.10	1.07	0.23	1.66	2.43
1	15	35	−0.80	1.04	−0.24	−0.63	−0.10	−0.90
1.5	15	35	1.63	0.38	−0.77	−1.25	−0.82	2.10
2	15	35	2.79	1.11	1.19	−1.08	0.53	3.91
5	15	35	−3.27	−15.86	1.10	−2.59	0.98	5.68
1	30	35	0.39	0.95	0.16	0.37	0.08	0.50
1.5	30	35	0.50	−0.00	0.76	−1.01	−0.67	0.81
2	30	35	−0.33	−1.33	1.22	−2.00	−0.07	0.52
5	30	35	−5.63	−18.94	1.29	−1.13	0.68	3.63
1	60	35	0.15	−0.81	0.30	0.14	0.34	0.16
1.5	60	35	−0.24	−1.88	1.38	−0.84	0.02	0.04
2	60	35	−0.00	−2.14	1.18	−0.48	−0.38	0.80
5	60	35	−7.86	−20.91	1.32	−0.87	0.05	1.43
1	15	50	−0.27	0.62	0.72	−0.32	−0.61	−0.33
1.5	15	50	1.46	0.25	1.26	−1.36	−0.39	1.75
2	15	50	2.62	−0.48	1.00	−1.38	0.32	3.57
5	15	50	−2.48	−15.50	0.34	−2.12	0.66	6.34
1	30	50	−0.27	−1.02	−0.19	−0.28	−0.25	−0.17
1.5	30	50	0.10	0.24	0.23	−1.43	−0.30	0.21
2	30	50	0.58	−1.93	0.32	−1.29	−0.14	1.55
5	30	50	−6.32	−19.04	0.90	−1.99	0.08	2.61
1	60	50	0.06	−0.31	0.07	0.06	0.32	0.13
1.5	60	50	0.11	1.24	0.47	−0.66	−0.18	0.30
2	60	50	0.33	−2.03	0.16	−0.22	0.34	1.33
5	60	50	−7.37	−21.35	0.35	−0.65	0.35	1.85

Table C.4: Bias ($\times 100$) in estimates of β from the simulations with $k = 3$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_{\cdot}$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	0.89	−0.26	−0.19	0.70	0.55	0.90
1.5	15	20	0.97	0.37	0.79	−0.71	−0.16	1.22
2	15	20	1.45	0.17	−0.16	−0.75	0.30	2.13
5	15	20	−2.85	−6.90	0.60	−1.88	−0.01	3.21
1	30	20	−0.00	0.58	0.06	−0.02	−0.14	0.01
1.5	30	20	0.49	0.86	−0.20	−0.43	−0.02	0.71
2	30	20	0.12	−0.62	0.48	−0.93	0.01	0.90
5	30	20	−4.45	−7.55	0.23	−1.04	0.38	2.06
1	60	20	−0.08	−0.17	0.21	−0.08	−0.05	−0.16
1.5	60	20	0.94	0.41	−0.10	0.50	0.92	1.18
2	60	20	−0.32	−1.37	0.59	−0.53	−0.23	0.54
5	60	20	−5.68	−9.82	0.70	−0.68	−0.18	1.10
1	15	35	0.45	0.33	0.57	0.40	0.33	0.43
1.5	15	35	0.59	0.82	0.48	−1.09	−0.26	0.75
2	15	35	1.72	0.31	−0.48	−0.61	0.62	2.35
5	15	35	−3.16	−6.17	0.03	−2.29	−0.37	3.01
1	30	35	−0.51	−0.43	−0.19	−0.53	−0.77	−0.59
1.5	30	35	0.36	0.01	−0.03	−0.69	−0.44	0.57
2	30	35	0.69	−0.86	0.26	−0.52	−0.01	1.36
5	30	35	−4.73	−8.18	−0.27	−1.29	−0.25	1.82
1	60	35	−0.26	−0.00	0.13	−0.23	−0.16	−0.31
1.5	60	35	0.36	−0.37	−0.68	−0.13	0.07	0.46
2	60	35	−0.33	−1.06	0.04	−0.69	−0.34	0.38
5	60	35	−5.28	−10.06	0.20	−0.10	0.55	1.65
1	15	50	0.26	−0.05	0.07	0.32	0.45	0.24
1.5	15	50	1.26	0.91	−0.18	−0.54	−0.01	1.43
2	15	50	1.19	0.67	−0.28	−1.26	−0.21	1.80
5	15	50	−3.27	−6.94	0.29	−2.36	−0.27	2.93
1	30	50	−0.16	−0.16	−0.17	−0.20	−0.30	−0.14
1.5	30	50	0.33	0.65	0.45	−0.69	−0.17	0.49
2	30	50	0.25	−0.78	0.63	−0.95	−0.09	0.90
5	30	50	−4.68	−8.41	0.06	−1.27	0.14	1.92
1	60	50	0.31	0.19	0.33	0.29	0.20	0.32
1.5	60	50	−0.23	0.06	0.09	−0.74	−0.06	−0.06
2	60	50	0.35	−0.66	−0.04	−0.01	0.39	1.09
5	60	50	−5.84	−9.98	−0.08	−0.71	0.32	1.05

Table C.5: Bias (×100) in estimates of β from the simulations with $k = 10$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	−0.04	0.12	−0.07	−0.01	−0.01	−0.03
1.5	15	20	0.81	0.78	−0.08	0.00	0.16	0.85
2	15	20	0.89	0.49	0.32	−0.33	−0.09	1.07
5	15	20	0.27	−0.78	−0.01	−0.27	0.25	2.22
1	30	20	−0.21	0.16	0.55	−0.19	−0.18	−0.21
1.5	30	20	0.35	0.52	0.17	−0.27	−0.06	0.40
2	30	20	0.63	0.47	0.51	−0.25	0.24	0.86
5	30	20	−0.77	−1.75	0.40	−0.50	0.23	1.46
1	60	20	−0.03	−0.18	−0.30	0.00	0.05	−0.03
1.5	60	20	0.51	−0.18	0.01	0.07	0.25	0.59
2	60	20	0.36	−0.30	0.10	−0.13	0.08	0.62
5	60	20	−1.28	−2.16	−0.46	−0.20	0.44	1.10
1	15	35	0.03	0.09	0.15	0.09	0.16	0.03
1.5	15	35	0.84	0.86	−0.42	−0.04	0.22	0.88
2	15	35	0.77	0.92	0.08	−0.50	−0.13	0.94
5	15	35	0.34	−0.61	−0.16	−0.31	0.26	2.37
1	30	35	0.15	0.19	0.12	0.17	0.24	0.15
1.5	30	35	0.37	−0.04	0.14	−0.30	0.01	0.42
2	30	35	0.68	0.50	−0.11	−0.24	0.19	0.89
5	30	35	−1.07	−1.95	0.31	−0.89	0.01	1.17
1	60	35	0.17	0.05	0.50	0.16	0.14	0.18
1.5	60	35	0.40	0.06	−0.04	−0.04	0.23	0.47
2	60	35	0.43	−0.31	0.14	−0.08	0.27	0.69
5	60	35	−1.65	−2.67	0.09	−0.52	0.10	0.81
1	15	50	0.06	−0.19	0.09	0.09	0.09	0.06
1.5	15	50	0.58	0.63	0.35	−0.39	−0.19	0.62
2	15	50	1.06	1.02	0.01	−0.29	0.12	1.23
5	15	50	−0.02	−1.05	−0.11	−0.71	−0.04	2.02
1	30	50	0.11	−0.23	0.07	0.10	0.09	0.11
1.5	30	50	0.55	0.43	−0.04	−0.16	0.13	0.59
2	30	50	0.83	0.45	0.14	−0.16	0.25	1.04
5	30	50	−0.74	−1.98	−0.06	−0.61	0.25	1.54
1	60	50	0.02	−0.35	0.09	−0.00	−0.06	0.02
1.5	60	50	0.03	−0.03	0.19	−0.42	−0.28	0.09
2	60	50	−0.08	0.08	0.17	−0.63	−0.35	0.17
5	60	50	−1.99	−2.47	−0.05	−0.89	−0.32	0.45

Table C.6: Standardized bias (100× bias/standard error) in estimates of β from the simulations with $k = 1$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	0.13	2.28	3.76	0.27	−1.38	0.36
1.5	15	20	3.67	6.47	−3.25	−6.63	−8.36	3.43
2	15	20	6.25	8.76	2.50	−7.37	−5.03	7.11
5	15	20	−8.61	−50.22	3.45	−13.67	−9.35	10.73
1	30	20	−2.94	0.24	2.14	−3.95	−2.28	−2.98
1.5	30	20	6.69	5.13	5.08	1.98	2.30	6.11
2	30	20	3.89	−3.05	3.74	−3.56	−6.29	4.28
5	30	20	−20.93	−58.50	6.32	−14.57	−10.88	1.42
1	60	20	−2.35	−2.90	−3.25	−3.08	0.97	−1.76
1.5	60	20	1.28	1.46	8.44	−2.19	−4.33	0.70
2	60	20	4.69	−3.40	8.43	2.04	0.01	5.19
5	60	20	−17.64	−70.66	5.67	1.42	−0.66	6.48
1	15	35	−2.25	3.70	−1.28	−2.70	−0.51	−2.48
1.5	15	35	6.38	5.35	−3.94	−6.22	−5.61	6.31
2	15	35	11.96	9.95	3.83	−4.78	−0.92	12.83
5	15	35	−8.46	−67.89	5.54	−12.27	−1.05	18.30
1	30	35	1.79	3.69	0.55	2.36	0.51	2.11
1.5	30	35	2.78	5.09	2.61	−5.27	−5.34	2.33
2	30	35	1.13	1.49	5.51	−9.80	−3.68	1.63
5	30	35	−18.36	−83.45	7.17	−5.72	−2.43	11.84
1	60	35	0.82	−3.07	1.06	1.04	2.11	0.68
1.5	60	35	0.85	−1.78	5.49	−3.56	−1.44	0.38
2	60	35	1.97	−1.12	4.82	−2.81	−5.79	2.24
5	60	35	−26.79	−94.37	7.85	−3.82	−5.05	5.01
1	15	50	−1.19	2.33	3.54	−1.89	−3.18	−1.33
1.5	15	50	6.89	4.87	5.46	−8.06	−4.28	6.63
2	15	50	11.86	3.36	4.55	−8.74	−1.23	12.97
5	15	50	−8.44	−79.86	2.70	−12.80	−1.57	23.72
1	30	50	−1.44	−3.90	−1.39	−2.10	−1.79	−1.02
1.5	30	50	1.13	5.23	0.04	−8.63	−3.52	0.38
2	30	50	3.99	−2.80	1.22	−7.83	−3.55	5.51
5	30	50	−26.11	−101.85	6.28	−11.10	−4.07	10.15
1	60	50	0.54	−1.46	−0.68	0.75	2.22	0.77
1.5	60	50	1.90	10.54	1.43	−3.43	−2.87	1.44
2	60	50	3.27	−2.93	1.10	−1.59	−1.06	5.18
5	60	50	−31.61	−115.96	3.28	−3.63	−2.56	7.42

Table C.7: Standardized bias (100× bias/standard error) in estimates of β from the simulations with $k = 3$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_\cdot$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	5.22	−0.55	−1.19	6.09	3.67	5.34
1.5	15	20	4.63	2.69	3.22	−5.59	−2.15	4.88
2	15	20	7.61	2.54	−2.75	−4.77	0.37	9.15
5	15	20	−13.99	−37.17	1.49	−12.20	−1.97	14.56
1	30	20	0.31	2.16	−0.22	0.68	0.06	0.40
1.5	30	20	3.43	5.70	−1.44	−2.72	−0.31	3.55
2	30	20	1.95	−0.65	1.81	−6.27	−0.76	4.26
5	30	20	−22.93	−42.90	1.90	−7.16	0.77	10.11
1	60	20	−0.85	−1.29	0.55	−1.03	−0.51	−1.18
1.5	60	20	5.23	4.17	−1.63	2.87	4.19	5.50
2	60	20	−0.56	−4.77	2.40	−4.33	−2.83	2.40
5	60	20	−30.18	−56.84	4.84	−4.49	−2.86	5.90
1	15	35	2.72	1.81	4.02	3.62	2.52	2.68
1.5	15	35	4.01	6.39	2.21	−9.62	−2.54	4.31
2	15	35	11.73	3.34	−4.00	−5.20	3.53	14.26
5	15	35	−20.70	−42.81	−1.06	−20.65	−3.39	18.52
1	30	35	−2.57	−2.73	−0.74	−4.04	−5.31	−3.03
1.5	30	35	2.53	1.14	−0.42	−7.01	−3.98	3.19
2	30	35	5.54	−4.84	1.38	−4.81	−0.81	8.64
5	30	35	−33.83	−59.60	−2.46	−13.28	−3.61	10.98
1	60	35	−1.76	−0.16	1.08	−2.17	−0.99	−2.05
1.5	60	35	3.20	−1.65	−5.34	−1.10	0.29	3.14
2	60	35	−1.02	−5.35	−0.27	−6.26	−3.37	2.72
5	60	35	−38.29	−75.88	1.44	−1.61	2.23	10.53
1	15	50	1.56	−0.25	1.04	3.14	3.80	1.43
1.5	15	50	9.59	7.98	−2.81	−5.70	−0.53	10.14
2	15	50	9.19	6.25	−3.53	−13.99	−2.96	12.38
5	15	50	−26.33	−58.84	1.61	−26.51	−3.90	21.10
1	30	50	−1.35	−1.24	−1.39	−2.40	−2.98	−1.19
1.5	30	50	2.79	5.59	3.59	−7.99	−2.44	3.33
2	30	50	2.40	−4.98	5.40	−11.27	−2.32	6.40
5	30	50	−39.84	−73.96	0.03	−14.78	−0.33	14.41
1	60	50	2.54	1.48	2.79	3.42	1.99	2.67
1.5	60	50	−1.03	2.34	0.16	−7.95	−1.00	−0.29
2	60	50	3.23	−4.15	−0.54	−1.02	2.35	7.99
5	60	50	−50.98	−89.54	−1.05	−8.56	0.84	7.90

Table C.8: Standardized bias (100× bias/standard error) in estimates of β from the simulations with $k = 10$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	−0.37	1.13	−0.56	−0.34	−0.19	−0.32
1.5	15	20	5.22	5.94	−1.23	−0.17	0.37	5.34
2	15	20	6.59	2.96	1.45	−2.88	−1.34	7.55
5	15	20	2.67	−6.49	−2.85	−0.94	1.71	17.46
1	30	20	−1.40	0.70	4.86	−1.81	−1.22	−1.37
1.5	30	20	2.70	4.47	0.48	−3.00	−0.65	2.94
2	30	20	5.31	4.45	4.29	−3.31	1.42	6.86
5	30	20	−7.07	−16.31	2.44	−5.69	1.33	12.09
1	60	20	−0.38	−1.60	−2.74	0.00	0.43	−0.32
1.5	60	20	4.82	−1.71	−0.13	1.07	2.10	5.26
2	60	20	3.35	−2.40	1.01	−2.01	0.42	5.34
5	60	20	−11.83	−20.82	−6.29	−2.19	3.74	10.37
1	15	35	0.86	0.89	1.31	1.61	2.04	0.92
1.5	15	35	8.27	8.44	−4.86	−0.30	1.80	8.55
2	15	35	7.70	9.69	0.17	−6.71	−1.93	9.14
5	15	35	3.66	−6.59	−3.70	−2.79	2.20	24.36
1	30	35	2.01	2.02	1.06	3.05	3.05	2.05
1.5	30	35	3.74	−0.47	1.75	−5.47	−0.68	4.08
2	30	35	7.58	5.73	−2.37	−4.10	1.54	9.67
5	30	35	−13.20	−23.96	3.28	−15.05	−0.79	12.75
1	60	35	1.98	0.47	6.47	2.71	1.81	2.08
1.5	60	35	4.63	1.06	−0.69	−1.18	2.35	5.25
2	60	35	5.01	−3.72	1.58	−2.17	2.70	7.81
5	60	35	−20.65	−33.98	0.59	−9.60	0.68	9.72
1	15	50	0.01	−2.25	1.20	0.80	0.57	0.02
1.5	15	50	6.86	6.98	3.36	−6.67	−2.74	7.16
2	15	50	13.02	12.66	−0.49	−4.25	1.48	14.85
5	15	50	−0.50	−14.10	−2.96	−11.64	−1.35	24.67
1	30	50	1.52	−3.05	1.01	1.86	1.22	1.50
1.5	30	50	7.47	5.82	−1.00	−2.65	1.85	7.95
2	30	50	11.10	5.95	1.59	−3.44	3.02	13.69
5	30	50	−10.02	−28.62	−1.83	−11.26	3.77	21.51
1	60	50	0.12	−5.00	1.38	−0.25	−0.95	0.12
1.5	60	50	0.57	−0.05	3.17	−8.55	−4.19	1.20
2	60	50	−0.57	1.47	2.15	−12.30	−5.08	2.86
5	60	50	−29.85	−37.48	−1.40	−18.77	−4.87	6.60

Table C.9: Mean square error (×100) in estimates of β from the simulations with $k = 1$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	17.12	17.29	13.20	13.67	12.80	18.09
1.5	15	20	17.09	14.61	11.68	14.29	12.67	18.43
2	15	20	15.23	13.36	11.95	13.37	12.82	17.09
5	15	20	11.67	11.96	8.90	12.83	12.19	16.28
1	30	20	15.98	15.10	12.16	14.13	12.35	16.97
1.5	30	20	15.14	13.45	11.89	13.79	12.22	16.47
2	30	20	15.29	12.61	10.20	14.72	12.57	17.26
5	30	20	12.48	12.23	8.13	14.35	12.27	16.81
1	60	20	15.43	15.41	12.01	14.45	12.12	16.71
1.5	60	20	14.72	13.71	11.72	14.23	12.01	16.35
2	60	20	14.62	12.84	10.67	14.81	12.34	16.54
5	60	20	11.62	12.67	8.03	14.29	11.79	15.87
1	15	35	9.28	9.99	6.93	7.37	7.33	9.77
1.5	15	35	9.29	8.54	6.80	7.58	6.88	9.91
2	15	35	9.29	7.65	5.90	8.09	7.43	10.28
5	15	35	7.27	7.57	4.80	8.18	7.21	10.26
1	30	35	9.31	8.76	6.82	8.07	6.98	9.72
1.5	30	35	9.10	8.29	6.61	8.26	7.32	9.72
2	30	35	8.66	7.31	5.94	8.24	7.21	9.64
5	30	35	6.94	8.65	4.63	8.19	7.05	9.58
1	60	35	9.16	8.24	6.58	8.50	7.19	9.64
1.5	60	35	8.69	7.73	6.10	8.30	7.27	9.41
2	60	35	8.42	7.38	5.66	8.39	7.30	9.56
5	60	35	7.12	9.31	4.34	8.52	7.02	9.62
1	15	50	6.79	6.91	4.92	5.38	5.22	7.07
1.5	15	50	6.70	6.21	4.18	5.55	5.16	7.03
2	15	50	6.66	5.54	4.34	5.77	5.19	7.45
5	15	50	5.07	5.95	3.49	5.80	5.11	7.35
1	30	50	6.35	6.13	4.59	5.51	4.90	6.48
1.5	30	50	6.27	5.61	4.23	5.68	4.97	6.67
2	30	50	6.40	5.14	4.04	6.04	5.08	7.15
5	30	50	5.03	7.23	3.25	5.76	5.09	6.56
1	60	50	6.04	5.85	4.18	5.60	4.97	6.38
1.5	60	50	5.92	5.54	4.23	5.63	4.99	6.48
2	60	50	5.87	5.19	3.80	5.90	5.02	6.58
5	60	50	4.95	8.10	3.22	5.78	4.79	6.43

Table C.10: Mean square error (×100) in estimates of β from the simulations with $k = 3$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	4.75	5.06	4.27	4.19	4.45	4.81
1.5	15	20	4.63	4.63	3.94	4.18	4.44	4.76
2	15	20	4.63	4.53	3.79	4.30	4.47	4.92
5	15	20	3.95	4.05	3.13	4.35	4.61	4.82
1	30	20	4.51	4.50	3.91	4.14	4.30	4.57
1.5	30	20	4.37	4.06	3.62	4.11	4.35	4.51
2	30	20	4.29	3.96	3.53	4.18	4.26	4.58
5	30	20	3.86	3.84	2.64	4.17	4.22	4.56
1	60	20	4.33	4.04	3.85	4.12	4.36	4.43
1.5	60	20	3.86	4.07	3.70	3.75	3.95	4.01
2	60	20	3.88	3.60	3.43	3.92	4.10	4.15
5	60	20	3.75	4.11	2.35	4.04	4.18	4.26
1	15	35	3.05	2.73	2.48	2.68	2.85	3.09
1.5	15	35	2.91	2.67	2.42	2.64	2.83	2.99
2	15	35	2.56	2.45	2.10	2.36	2.50	2.71
5	15	35	2.26	2.40	1.76	2.42	2.49	2.74
1	30	35	2.53	2.56	2.12	2.31	2.47	2.58
1.5	30	35	2.53	2.44	2.07	2.37	2.49	2.60
2	30	35	2.30	2.08	1.96	2.24	2.47	2.48
5	30	35	2.35	2.58	1.54	2.49	2.58	2.70
1	60	35	2.44	2.27	2.09	2.32	2.42	2.49
1.5	60	35	2.39	2.24	2.01	2.33	2.52	2.48
2	60	35	2.09	2.09	1.92	2.12	2.31	2.24
5	60	35	2.03	2.92	1.48	2.10	2.29	2.24
1	15	50	2.01	1.98	1.80	1.76	1.88	2.03
1.5	15	50	1.94	1.93	1.56	1.74	1.87	1.98
2	15	50	1.90	1.73	1.57	1.74	1.83	2.02
5	15	50	1.57	1.90	1.21	1.68	1.72	1.90
1	30	50	1.76	1.79	1.54	1.61	1.74	1.80
1.5	30	50	1.78	1.62	1.44	1.66	1.81	1.83
2	30	50	1.67	1.51	1.37	1.64	1.77	1.77
5	30	50	1.59	2.10	1.06	1.61	1.75	1.75
1	60	50	1.66	1.67	1.43	1.58	1.69	1.68
1.5	60	50	1.66	1.51	1.31	1.62	1.70	1.71
2	60	50	1.61	1.51	1.30	1.63	1.70	1.72
5	60	50	1.61	2.27	1.01	1.52	1.63	1.62

Table C.11: Mean square error (×100) in estimates of β from the simulations with $k = 10$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	1.95	1.94	1.82	1.87	1.91	1.95
1.5	15	20	1.81	1.76	1.65	1.74	1.79	1.83
2	15	20	1.74	1.71	1.55	1.72	1.75	1.77
5	15	20	1.44	1.51	1.22	1.54	1.56	1.57
1	30	20	1.45	1.58	1.40	1.42	1.52	1.45
1.5	30	20	1.40	1.41	1.34	1.36	1.46	1.41
2	30	20	1.32	1.34	1.30	1.30	1.40	1.35
5	30	20	1.17	1.30	1.02	1.23	1.35	1.27
1	60	20	1.33	1.28	1.22	1.29	1.41	1.33
1.5	60	20	1.28	1.21	1.10	1.25	1.40	1.29
2	60	20	1.22	1.14	1.09	1.20	1.31	1.25
5	60	20	1.16	1.13	0.85	1.20	1.33	1.24
1	15	35	1.04	1.13	0.97	1.00	1.04	1.04
1.5	15	35	1.02	1.08	0.97	0.98	1.02	1.03
2	15	35	0.92	0.96	0.91	0.89	0.93	0.94
5	15	35	0.84	0.84	0.71	0.90	0.92	0.96
1	30	35	0.88	0.88	0.77	0.84	0.90	0.88
1.5	30	35	0.81	0.77	0.76	0.79	0.87	0.81
2	30	35	0.78	0.80	0.71	0.75	0.81	0.80
5	30	35	0.74	0.72	0.54	0.76	0.82	0.80
1	60	35	0.75	0.71	0.69	0.73	0.83	0.75
1.5	60	35	0.65	0.71	0.67	0.64	0.72	0.66
2	60	35	0.67	0.69	0.62	0.67	0.75	0.69
5	60	35	0.66	0.71	0.48	0.67	0.74	0.69
1	15	50	0.76	0.72	0.72	0.72	0.74	0.76
1.5	15	50	0.72	0.71	0.68	0.69	0.72	0.73
2	15	50	0.68	0.66	0.57	0.65	0.67	0.69
5	15	50	0.58	0.61	0.50	0.63	0.62	0.66
1	30	50	0.60	0.61	0.54	0.58	0.63	0.60
1.5	30	50	0.57	0.53	0.54	0.55	0.61	0.57
2	30	50	0.58	0.55	0.52	0.57	0.63	0.60
5	30	50	0.52	0.50	0.39	0.53	0.58	0.57
1	60	50	0.51	0.48	0.48	0.49	0.54	0.50
1.5	60	50	0.46	0.50	0.47	0.45	0.51	0.47
2	60	50	0.47	0.45	0.45	0.47	0.54	0.48
5	60	50	0.48	0.53	0.36	0.47	0.54	0.48

Table C.12: L-lower Percent coverage of 95% confidence limit for β, based on the simulations with $k = 1$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_{.}$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	3.55	5.05	4.55	8.80	6.80	3.55
1.5	15	20	4.05	4.95	3.90	7.80	6.25	3.90
2	15	20	3.80	5.10	4.75	7.20	5.60	3.65
5	15	20	2.80	0.75	4.50	6.60	3.35	3.35
1	30	20	3.70	4.20	4.00	7.65	7.15	3.45
1.5	30	20	4.30	4.45	3.60	8.35	7.10	3.70
2	30	20	3.80	3.15	3.90	9.25	5.45	4.05
5	30	20	2.35	1.15	3.80	6.40	3.45	3.35
1	60	20	3.20	4.15	3.85	8.30	7.85	3.30
1.5	60	20	3.60	4.25	4.80	7.60	5.65	3.60
2	60	20	4.30	4.40	4.65	8.95	6.25	3.60
5	60	20	2.55	0.25	4.60	8.40	3.15	3.50
1	15	35	2.80	4.85	3.15	7.95	7.20	2.80
1.5	15	35	3.15	3.95	3.20	6.15	5.60	2.95
2	15	35	5.05	4.25	2.90	8.05	7.20	4.95
5	15	35	2.85	0.55	3.20	7.05	4.15	4.05
1	30	35	2.70	3.55	3.20	8.35	6.95	3.35
1.5	30	35	3.75	3.65	4.10	7.70	7.30	3.35
2	30	35	3.20	3.65	3.65	6.95	6.05	3.00
5	30	35	2.00	0.25	3.75	6.80	3.70	3.25
1	60	35	3.60	3.00	3.40	9.80	7.55	3.60
1.5	60	35	2.90	3.60	4.05	6.90	6.70	3.05
2	60	35	3.30	4.00	3.85	8.15	6.35	3.35
5	60	35	2.00	0.30	3.65	8.10	3.80	3.60
1	15	50	2.80	3.70	3.85	7.30	7.75	2.95
1.5	15	50	3.80	3.30	2.80	7.40	7.00	3.70
2	15	50	3.70	3.05	3.35	6.45	5.90	3.40
5	15	50	2.65	0.45	3.30	6.95	3.60	4.85
1	30	50	2.75	2.85	2.95	7.15	7.15	2.50
1.5	30	50	3.10	4.20	2.70	6.95	6.45	2.65
2	30	50	3.45	2.85	3.20	7.95	5.50	3.30
5	30	50	1.50	0.25	3.00	6.65	3.25	3.60
1	60	50	2.60	2.80	2.90	7.85	8.25	2.40
1.5	60	50	3.45	4.70	3.55	7.85	7.10	3.05
2	60	50	3.10	3.00	2.90	7.25	7.25	2.85
5	60	50	1.35	0.10	3.65	7.15	3.60	3.60

Table C.13: L-lower Percent coverage of 95% confidence limit for β, based on the simulations with $k = 3$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	μ.	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	3.85	3.60	3.15	9.60	4.90	3.75
1.5	15	20	4.15	3.40	3.55	8.50	4.95	4.20
2	15	20	4.30	4.15	3.00	9.20	4.70	4.35
5	15	20	2.90	2.00	3.50	9.30	4.95	4.85
1	30	20	3.75	3.70	3.75	9.85	5.45	3.80
1.5	30	20	4.30	3.30	3.65	9.60	5.80	4.45
2	30	20	4.45	4.20	3.60	9.05	5.45	4.50
5	30	20	2.70	1.55	3.90	8.80	4.45	4.70
1	60	20	3.80	4.00	4.05	9.50	6.00	3.75
1.5	60	20	3.30	4.45	3.30	8.60	5.10	3.25
2	60	20	3.95	3.10	4.25	8.35	4.90	4.05
5	60	20	2.20	1.15	3.35	8.70	4.40	3.95
1	15	35	3.15	2.90	3.15	9.35	4.70	3.25
1.5	15	35	3.60	3.85	3.60	8.10	5.40	3.65
2	15	35	4.10	3.40	2.85	8.15	4.35	4.05
5	15	35	2.35	1.15	3.10	7.05	3.90	4.50
1	30	35	2.95	2.80	2.85	8.25	5.15	3.05
1.5	30	35	3.55	3.45	3.15	8.70	5.30	3.50
2	30	35	3.55	1.90	2.90	7.45	4.30	3.65
5	30	35	1.85	1.05	3.10	7.65	4.30	4.20
1	60	35	3.30	2.55	3.75	8.80	5.25	3.30
1.5	60	35	3.65	2.95	2.65	8.25	5.65	3.55
2	60	35	2.85	2.80	3.40	7.40	4.70	3.10
5	60	35	1.25	0.90	3.20	7.75	3.70	3.40
1	15	50	2.30	2.85	3.00	8.10	4.50	2.50
1.5	15	50	3.50	3.65	2.80	7.85	4.00	3.50
2	15	50	4.50	3.80	3.25	7.40	4.35	4.65
5	15	50	1.70	0.75	3.10	6.20	3.35	4.40
1	30	50	3.55	3.15	3.00	7.95	5.35	3.70
1.5	30	50	3.25	3.75	3.25	8.20	5.05	3.20
2	30	50	3.10	2.00	3.10	7.75	5.15	3.30
5	30	50	1.25	0.80	2.60	6.80	4.55	3.65
1	60	50	2.80	3.20	3.20	8.75	5.20	3.00
1.5	60	50	3.40	2.65	2.45	7.55	4.90	3.50
2	60	50	3.05	2.55	3.10	8.80	5.10	3.20
5	60	50	0.85	0.20	2.80	7.20	3.45	3.20

Table C.14: L-lower Percent coverage of 95% confidence limit for β, based on the simulations with $k = 10$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	4.20	3.75	3.70	9.15	4.60	4.25
1.5	15	20	3.90	4.20	3.00	9.81	4.25	4.10
2	15	20	3.80	3.85	3.85	10.30	4.15	3.90
5	15	20	3.65	3.65	3.10	12.60	3.90	4.55
1	30	20	3.65	3.85	4.25	9.35	4.15	3.65
1.5	30	20	3.55	4.00	3.55	9.00	4.45	3.40
2	30	20	3.95	4.20	4.30	9.70	4.10	4.10
5	30	20	2.65	3.15	3.90	10.15	4.00	4.00
1	60	20	3.80	3.65	3.60	10.25	5.10	3.80
1.5	60	20	3.80	3.15	3.00	10.51	5.20	3.90
2	60	20	3.55	3.15	3.75	9.80	4.50	3.90
5	60	20	3.30	2.50	3.00	10.30	4.50	4.45
1	15	35	3.35	2.95	3.10	8.36	4.15	3.40
1.5	15	35	3.55	4.20	2.85	8.85	4.40	3.55
2	15	35	3.70	3.75	3.70	8.80	3.55	3.70
5	15	35	3.80	3.00	2.60	11.86	4.35	5.25
1	30	35	3.40	3.45	2.50	9.65	4.15	3.40
1.5	30	35	3.30	2.25	3.65	8.85	4.15	3.25
2	30	35	3.55	3.60	3.00	8.65	4.10	3.75
5	30	35	3.00	1.80	2.85	8.50	4.50	4.55
1	60	35	2.85	2.50	3.95	9.65	5.00	2.90
1.5	60	35	2.65	3.40	3.05	7.95	3.30	2.60
2	60	35	2.75	3.00	3.40	8.60	4.35	2.80
5	60	35	2.35	1.55	2.75	8.00	4.30	4.30
1	15	50	3.10	2.40	3.25	8.90	3.75	3.15
1.5	15	50	3.05	3.30	3.10	8.15	3.50	3.05
2	15	50	3.75	3.50	2.45	9.10	3.95	3.75
5	15	50	2.35	2.25	2.90	10.60	2.90	4.65
1	30	50	3.15	2.85	2.90	9.65	4.75	3.20
1.5	30	50	4.00	2.85	2.95	9.00	4.90	3.95
2	30	50	3.50	3.40	3.55	8.60	4.60	3.65
5	30	50	2.60	1.40	2.85	8.70	4.50	4.50
1	60	50	2.75	2.45	2.70	8.80	3.55	2.70
1.5	60	50	2.95	2.95	3.10	7.60	4.05	3.00
2	60	50	2.75	2.45	3.55	7.35	4.25	2.85
5	60	50	1.50	2.00	2.70	6.45	3.35	3.35

Table C.15: U-Upper Percent coverage of 95% confidence limit for β, based on the simulations with $k = 1$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	3.50	4.20	3.25	8.80	6.80	3.45
1.5	15	20	3.15	2.95	3.85	9.70	8.10	3.25
2	15	20	2.80	2.30	3.05	8.90	7.50	3.00
5	15	20	4.15	6.70	3.05	9.35	6.35	3.00
1	30	20	3.75	3.95	3.60	10.00	8.10	3.65
1.5	30	20	3.10	3.05	3.65	7.50	6.60	2.90
2	30	20	3.20	3.35	2.80	9.75	8.15	3.15
5	30	20	4.95	7.40	2.55	10.20	6.80	3.65
1	60	20	3.55	4.85	3.95	8.55	6.65	3.95
1.5	60	20	2.80	3.45	3.90	8.95	7.45	2.75
2	60	20	2.75	2.90	3.15	8.35	7.40	2.80
5	60	20	4.45	10.20	3.00	7.55	5.25	3.00
1	15	35	2.85	3.00	3.25	8.30	7.35	3.40
1.5	15	35	2.25	2.75	3.85	8.90	7.30	2.50
2	15	35	2.80	2.60	2.45	9.50	7.55	2.60
5	15	35	3.10	9.50	2.60	9.85	5.45	2.10
1	30	35	3.00	3.00	3.05	8.30	7.25	2.80
1.5	30	35	2.90	2.35	3.35	8.65	7.50	2.70
2	30	35	3.00	2.30	1.95	9.35	7.85	3.35
5	30	35	4.35	12.25	1.85	9.05	5.85	2.90
1	60	35	3.35	3.65	3.25	8.20	7.55	3.15
1.5	60	35	3.05	2.85	2.40	8.85	7.75	2.60
2	60	35	3.20	3.20	2.65	8.30	7.65	3.50
5	60	35	4.90	14.90	2.30	7.95	5.55	2.80
1	15	50	2.80	3.00	3.05	7.80	7.80	2.90
1.5	15	50	2.45	2.65	2.00	9.10	7.70	2.60
2	15	50	2.60	2.60	2.60	10.20	7.30	3.00
5	15	50	3.45	11.80	2.40	9.85	5.55	1.40
1	30	50	2.80	2.95	2.95	8.00	7.60	2.80
1.5	30	50	3.55	2.10	2.85	8.90	7.85	3.30
2	30	50	3.05	2.60	2.70	9.40	7.70	2.85
5	30	50	4.65	16.70	2.25	9.25	5.90	2.60
1	60	50	2.75	3.00	2.65	7.65	7.10	2.75
1.5	60	50	2.85	2.05	2.70	8.50	8.40	3.15
2	60	50	3.00	2.50	2.70	8.50	7.70	3.20
5	60	50	5.00	21.30	2.15	8.15	5.90	2.25

Table C.16: U-Upper Percent coverage of 95% confidence limit for β, based on the simulations with $k = 3$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	2.80	3.55	3.75	8.10	4.10	2.65
1.5	15	20	2.40	3.25	3.40	9.60	4.85	2.55
2	15	20	2.75	3.20	4.25	10.00	4.70	2.85
5	15	20	3.90	5.80	3.30	12.55	5.35	2.45
1	30	20	3.60	3.40	3.20	10.25	5.30	3.55
1.5	30	20	3.35	2.70	3.50	9.60	5.25	3.20
2	30	20	3.15	3.45	4.30	9.90	4.85	3.20
5	30	20	5.00	7.65	3.05	10.20	4.95	2.75
1	60	20	3.65	3.75	3.95	8.85	5.40	3.85
1.5	60	20	3.30	3.40	3.95	8.10	4.85	3.30
2	60	20	3.40	3.35	4.05	9.85	5.45	3.45
5	60	20	5.80	8.70	2.35	9.55	5.30	3.70
1	15	35	3.80	2.35	2.85	9.10	5.40	3.90
1.5	15	35	3.55	2.85	3.25	10.55	6.30	3.70
2	15	35	1.90	3.15	2.75	9.30	4.25	1.85
5	15	35	4.35	6.80	2.90	11.15	4.20	1.75
1	30	35	3.10	3.60	2.90	8.85	5.30	3.10
1.5	30	35	3.00	2.80	3.05	10.15	5.80	3.05
2	30	35	2.20	3.10	3.05	7.95	5.05	2.30
5	30	35	6.00	8.30	2.60	11.50	5.45	2.90
1	60	35	3.40	2.50	2.85	9.25	5.75	3.50
1.5	60	35	3.10	2.90	3.20	8.75	5.70	3.05
2	60	35	2.80	3.30	3.35	8.55	5.25	2.70
5	60	35	4.80	13.45	2.60	7.50	3.60	1.90
1	15	50	3.50	3.05	3.30	7.80	5.10	3.70
1.5	15	50	2.65	2.35	3.30	9.35	4.90	2.80
2	15	50	2.05	2.15	3.65	10.15	5.10	2.05
5	15	50	4.75	8.60	3.25	12.15	4.25	1.25
1	30	50	2.65	2.75	3.50	7.55	4.80	2.75
1.5	30	50	2.65	2.20	2.55	9.30	5.90	2.80
2	30	50	2.60	2.70	2.10	10.60	4.90	2.70
5	30	50	6.50	12.10	3.05	9.85	4.25	1.95
1	60	50	2.50	3.05	2.80	7.00	5.35	2.65
1.5	60	50	2.75	2.20	2.20	9.70	5.40	2.55
2	60	50	2.85	3.20	2.40	8.90	5.45	2.50
5	60	50	7.25	14.30	2.55	8.90	4.35	2.55

Table C.17: U-Upper Percent coverage of 95% confidence limit for β, based on the simulations with $k = 10$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	3.80	3.20	3.40	9.25	4.95	3.95
1.5	15	20	3.30	3.15	3.95	9.36	4.15	3.30
2	15	20	3.50	3.95	4.30	10.30	4.85	3.50
5	15	20	3.35	4.10	3.70	12.25	4.20	2.45
1	30	20	3.35	4.90	3.25	9.50	4.55	3.35
1.5	30	20	2.95	3.15	3.90	10.80	4.60	3.00
2	30	20	2.50	3.20	3.40	10.30	3.80	2.45
5	30	20	3.45	4.45	4.00	10.65	3.65	2.40
1	60	20	3.95	3.45	3.55	9.95	4.50	4.05
1.5	60	20	2.80	3.80	3.25	9.75	4.65	2.90
2	60	20	3.50	3.60	3.55	10.10	4.35	3.45
5	60	20	4.30	4.70	4.30	10.50	4.30	2.90
1	15	35	2.65	2.95	2.60	8.41	3.50	2.75
1.5	15	35	2.85	3.15	3.50	8.75	4.15	2.85
2	15	35	3.00	2.55	3.20	10.05	3.95	2.95
5	15	35	2.35	3.60	3.30	12.56	3.85	1.50
1	30	35	3.35	2.60	3.50	9.20	3.75	3.40
1.5	30	35	2.75	2.75	3.00	10.00	4.35	2.60
2	30	35	2.75	2.95	3.40	9.65	4.50	2.65
5	30	35	4.80	4.80	2.90	11.90	4.45	2.40
1	60	35	3.50	2.85	2.75	8.60	4.80	3.45
1.5	60	35	2.25	2.75	3.15	8.05	3.95	2.20
2	60	35	2.40	3.25	3.30	9.20	4.30	2.40
5	60	35	4.55	6.15	2.65	10.25	4.00	2.70
1	15	50	2.90	2.90	3.05	8.70	3.80	3.05
1.5	15	50	2.70	2.65	2.55	10.45	4.55	2.75
2	15	50	2.15	2.40	2.55	9.75	3.60	2.10
5	15	50	2.85	4.40	3.30	13.70	3.95	1.80
1	30	50	2.90	3.35	2.20	8.65	4.00	2.90
1.5	30	50	2.15	2.35	3.05	8.40	3.65	2.15
2	30	50	2.85	2.60	2.80	9.45	4.35	2.70
5	30	50	4.00	4.05	3.60	10.90	4.10	1.90
1	60	50	3.55	2.75	3.35	8.20	4.55	3.30
1.5	60	50	2.40	2.95	2.65	8.90	4.10	2.40
2	60	50	3.00	2.35	2.80	9.95	4.65	2.70
5	60	50	5.30	6.25	3.30	11.40	4.50	2.80

Table C.18: T-Total Percent coverage of 95% confidence limits for β, based on the simulations with $k = 1$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	7.05	9.25	7.80	17.60	13.60	7.00
1.5	15	20	7.20	7.90	7.75	17.50	14.35	7.15
2	15	20	6.60	7.40	7.80	16.10	13.10	6.65
5	15	20	6.95	7.45	7.55	15.95	9.70	6.35
1	30	20	7.45	8.15	7.60	17.65	15.25	7.10
1.5	30	20	7.40	7.50	7.25	15.85	13.70	6.60
2	30	20	7.00	6.50	6.70	19.00	13.60	7.20
5	30	20	7.30	8.55	6.35	16.60	10.25	7.00
1	60	20	6.75	9.00	7.80	16.85	14.50	7.25
1.5	60	20	6.40	7.70	8.70	16.55	13.10	6.35
2	60	20	7.05	7.30	7.80	17.30	13.65	6.40
5	60	20	7.00	10.45	7.60	15.95	8.40	6.50
1	15	35	5.65	7.85	6.40	16.25	14.55	6.20
1.5	15	35	5.40	6.70	7.05	15.05	12.90	5.45
2	15	35	7.85	6.85	5.35	17.55	14.75	7.55
5	15	35	5.95	10.05	5.80	16.90	9.60	6.15
1	30	35	5.70	6.55	6.25	16.65	14.20	6.15
1.5	30	35	6.65	6.00	7.45	16.35	14.80	6.05
2	30	35	6.20	5.95	5.60	16.30	13.90	6.35
5	30	35	6.35	12.50	5.60	15.85	9.55	6.15
1	60	35	6.95	6.65	6.65	18.00	15.10	6.75
1.5	60	35	5.95	6.45	6.45	15.75	14.45	5.65
2	60	35	6.50	7.20	6.50	16.45	14.00	6.85
5	60	35	6.90	15.20	5.95	16.05	9.35	6.40
1	15	50	5.60	6.70	6.90	15.10	15.55	5.85
1.5	15	50	6.25	5.95	4.80	16.50	14.70	6.30
2	15	50	6.30	5.65	5.95	16.65	13.20	6.40
5	15	50	6.10	12.25	5.70	16.80	9.15	6.25
1	30	50	5.55	5.80	5.90	15.15	14.75	5.30
1.5	30	50	6.65	6.30	5.55	15.85	14.30	5.95
2	30	50	6.50	5.45	5.90	17.35	13.20	6.15
5	30	50	6.15	16.95	5.25	15.90	9.15	6.20
1	60	50	5.35	5.80	5.55	15.50	15.35	5.15
1.5	60	50	6.30	6.75	6.25	16.35	15.50	6.20
2	60	50	6.10	5.50	5.60	15.75	14.95	6.05
5	60	50	6.35	21.40	5.80	15.30	9.50	5.85

Table C.19: T-Total Percent coverage of 95% confidence limits for β, based on the simulations with $k = 3$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	6.65	7.15	6.90	17.70	9.00	6.40
1.5	15	20	6.55	6.65	6.95	18.10	9.80	6.75
2	15	20	7.05	7.35	7.25	19.20	9.40	7.20
5	15	20	6.80	7.80	6.80	21.85	10.30	7.30
1	30	20	7.35	7.10	6.95	20.10	10.75	7.35
1.5	30	20	7.65	6.00	7.15	19.20	11.05	7.65
2	30	20	7.60	7.65	7.90	18.95	10.30	7.70
5	30	20	7.70	9.20	6.95	19.00	9.40	7.45
1	60	20	7.45	7.75	8.00	18.35	11.40	7.60
1.5	60	20	6.60	7.85	7.25	16.70	9.95	6.55
2	60	20	7.35	6.45	8.30	18.20	10.35	7.50
5	60	20	8.00	9.85	5.70	18.25	9.70	7.65
1	15	35	6.95	5.25	6.00	18.45	10.10	7.15
1.5	15	35	7.15	6.70	6.85	18.65	11.70	7.35
2	15	35	6.00	6.55	5.60	17.45	8.60	5.90
5	15	35	6.70	7.95	6.00	18.20	8.10	6.25
1	30	35	6.05	6.40	5.75	17.10	10.45	6.15
1.5	30	35	6.55	6.25	6.20	18.85	11.10	6.55
2	30	35	5.75	5.00	5.95	15.40	9.35	5.95
5	30	35	7.85	9.35	5.70	19.15	9.75	7.10
1	60	35	6.70	5.05	6.60	18.05	11.00	6.80
1.5	60	35	6.75	5.85	5.85	17.00	11.35	6.60
2	60	35	5.65	6.10	6.75	15.95	9.95	5.80
5	60	35	6.05	14.35	5.80	15.25	7.30	5.30
1	15	50	5.80	5.90	6.30	15.90	9.60	6.20
1.5	15	50	6.15	6.00	6.10	17.20	8.90	6.30
2	15	50	6.55	5.95	6.90	17.55	9.45	6.70
5	15	50	6.45	9.35	6.35	18.35	7.60	5.65
1	30	50	6.20	5.90	6.50	15.50	10.15	6.45
1.5	30	50	5.90	5.95	5.80	17.50	10.95	6.00
2	30	50	5.70	4.70	5.20	18.35	10.05	6.00
5	30	50	7.75	12.90	5.65	16.65	8.80	5.60
1	60	50	5.30	6.25	6.00	15.75	10.55	5.65
1.5	60	50	6.15	4.85	4.65	17.25	10.30	6.05
2	60	50	5.90	5.75	5.50	17.70	10.55	5.70
5	60	50	8.10	14.50	5.35	16.10	7.80	5.75

Table C.20: T-Total Percent coverage of 95% confidence limits for β, based on the simulations with $k = 10$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	8.00	6.95	7.10	18.40	9.55	8.20
1.5	15	20	7.21	7.35	6.95	19.17	8.41	7.41
2	15	20	7.30	7.80	8.15	20.60	9.00	7.40
5	15	20	7.00	7.75	6.80	24.85	8.10	7.00
1	30	20	7.00	8.75	7.50	18.85	8.70	7.00
1.5	30	20	6.50	7.15	7.45	19.80	9.05	6.40
2	30	20	6.45	7.40	7.70	20.00	7.90	6.55
5	30	20	6.10	7.60	7.90	20.80	7.65	6.40
1	60	20	7.75	7.10	7.15	20.20	9.60	7.85
1.5	60	20	6.60	6.95	6.25	20.26	9.85	6.80
2	60	20	7.05	6.75	7.30	19.90	8.85	7.35
5	60	20	7.60	7.20	7.30	20.80	8.80	7.35
1	15	35	6.01	5.90	5.70	16.77	7.66	6.16
1.5	15	35	6.40	7.35	6.35	17.60	8.55	6.40
2	15	35	6.70	6.30	6.90	18.85	7.50	6.65
5	15	35	6.15	6.60	5.90	24.41	8.20	6.75
1	30	35	6.75	6.05	6.00	18.85	7.90	6.80
1.5	30	35	6.05	5.00	6.65	18.85	8.50	5.85
2	30	35	6.30	6.55	6.40	18.30	8.60	6.40
5	30	35	7.80	6.60	5.75	20.40	8.95	6.95
1	60	35	6.35	5.35	6.70	18.25	9.80	6.35
1.5	60	35	4.90	6.15	6.20	16.00	7.25	4.80
2	60	35	5.15	6.25	6.70	17.80	8.65	5.20
5	60	35	6.90	7.70	5.40	18.25	8.30	7.00
1	15	50	6.00	5.30	6.30	17.61	7.55	6.20
1.5	15	50	5.75	5.95	5.65	18.60	8.05	5.80
2	15	50	5.90	5.90	5.00	18.86	7.55	5.85
5	15	50	5.20	6.65	6.20	24.30	6.85	6.45
1	30	50	6.05	6.20	5.10	18.30	8.75	6.10
1.5	30	50	6.15	5.20	6.00	17.40	8.55	6.10
2	30	50	6.35	6.00	6.35	18.05	8.95	6.35
5	30	50	6.60	5.45	6.45	19.60	8.60	6.40
1	60	50	6.30	5.20	6.05	17.00	8.10	6.00
1.5	60	50	5.35	5.90	5.75	16.50	8.15	5.40
2	60	50	5.75	4.80	6.35	17.30	8.90	5.55
5	60	50	6.80	8.25	6.00	17.85	7.85	6.15

Table C.21: L-lower Percent coverage of 90% confidence limit for β, based on the simulations with $k = 1$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	6.50	8.40	7.30	11.65	10.65	6.70
1.5	15	20	6.75	8.10	6.10	11.75	9.25	6.40
2	15	20	6.40	7.90	7.40	10.40	10.05	6.45
5	15	20	4.50	1.90	7.00	9.50	5.95	6.45
1	30	20	5.75	6.90	7.15	11.50	10.00	5.55
1.5	30	20	6.70	6.60	6.30	12.80	10.70	6.10
2	30	20	7.35	6.00	6.75	12.35	9.15	7.30
5	30	20	4.40	2.35	6.85	9.95	5.95	5.75
1	60	20	6.15	7.20	6.45	12.00	11.65	5.90
1.5	60	20	6.20	7.00	8.05	11.25	9.80	5.80
2	60	20	7.05	6.75	8.10	12.75	9.95	7.05
5	60	20	4.10	0.90	7.80	11.75	5.75	6.20
1	15	35	5.55	7.35	5.60	11.05	10.55	5.60
1.5	15	35	5.70	6.50	5.35	10.65	9.25	5.65
2	15	35	7.70	7.10	6.05	11.45	11.00	7.85
5	15	35	5.00	0.90	5.75	10.90	6.90	8.05
1	30	35	5.65	6.50	6.30	12.45	10.65	5.70
1.5	30	35	6.50	6.85	6.90	11.05	10.70	5.90
2	30	35	5.70	7.20	6.90	10.45	9.90	5.55
5	30	35	3.55	0.80	6.50	11.30	6.35	6.80
1	60	35	6.80	4.90	7.00	12.75	12.20	6.60
1.5	60	35	5.70	6.10	6.05	10.90	9.95	5.85
2	60	35	6.15	6.55	6.55	11.40	10.05	5.95
5	60	35	3.55	0.85	6.60	11.90	6.70	6.60
1	15	50	4.45	7.00	6.55	11.25	11.15	4.70
1.5	15	50	6.80	6.95	5.60	10.90	9.90	6.60
2	15	50	6.70	5.65	5.90	10.00	9.65	6.70
5	15	50	5.15	0.65	6.25	10.45	6.45	8.85
1	30	50	4.90	5.40	5.70	10.85	10.55	4.80
1.5	30	50	5.80	6.75	5.50	10.15	10.15	5.65
2	30	50	6.95	6.35	5.55	11.80	9.20	6.60
5	30	50	3.60	0.60	6.05	9.70	6.45	6.05
1	60	50	5.30	5.35	5.45	11.50	12.55	5.40
1.5	60	50	6.15	7.35	6.25	10.10	10.80	5.90
2	60	50	5.35	6.05	5.50	10.65	11.05	5.50
5	60	50	3.30	0.25	6.25	11.05	6.40	6.45

Table C.22: L-lower Percent coverage of 90% confidence limit for β, based on the simulations with $k = 3$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	6.45	6.50	5.75	13.85	8.45	6.50
1.5	15	20	6.90	6.10	5.90	12.00	7.80	6.60
2	15	20	7.25	6.75	5.55	13.10	8.20	7.30
5	15	20	5.30	3.15	6.30	12.20	7.65	8.45
1	30	20	7.00	6.75	6.25	13.85	8.95	6.95
1.5	30	20	7.55	7.05	6.00	13.15	9.40	7.50
2	30	20	7.10	7.60	5.60	12.90	8.05	7.60
5	30	20	4.80	2.90	6.30	13.30	7.60	7.95
1	60	20	6.50	6.45	6.40	13.60	9.00	6.40
1.5	60	20	6.30	7.05	6.00	12.45	8.95	6.20
2	60	20	6.80	5.45	6.40	12.50	7.80	6.80
5	60	20	3.95	2.20	5.75	11.75	6.85	7.20
1	15	35	6.40	5.10	6.50	13.55	8.80	6.50
1.5	15	35	6.40	6.60	6.70	11.90	8.55	6.45
2	15	35	7.35	6.35	5.35	11.20	7.80	7.30
5	15	35	4.10	2.35	5.90	10.55	6.85	7.70
1	30	35	5.85	5.05	5.20	11.45	8.10	5.70
1.5	30	35	6.55	6.00	5.80	12.60	8.45	6.70
2	30	35	5.80	3.90	5.75	11.20	7.55	6.05
5	30	35	3.55	1.95	5.15	11.90	7.05	7.30
1	60	35	5.70	5.15	6.45	12.05	8.75	6.00
1.5	60	35	6.55	5.85	4.95	12.65	8.85	6.25
2	60	35	5.30	5.55	5.45	11.00	7.85	5.65
5	60	35	2.60	1.50	5.55	11.70	7.30	6.30
1	15	50	5.10	5.50	5.20	12.85	8.65	5.55
1.5	15	50	6.70	6.90	5.15	11.60	7.85	6.75
2	15	50	7.45	6.50	5.35	10.70	8.10	7.65
5	15	50	3.15	1.70	5.30	9.30	6.05	7.70
1	30	50	5.85	5.50	4.90	11.40	8.30	5.95
1.5	30	50	6.25	5.75	6.00	11.20	8.25	6.40
2	30	50	6.40	4.25	5.25	11.25	7.95	6.80
5	30	50	2.60	1.30	4.90	10.50	7.85	7.05
1	60	50	5.45	5.55	5.90	12.70	8.50	5.55
1.5	60	50	5.75	5.65	5.00	10.55	7.70	5.60
2	60	50	6.00	5.20	5.60	13.35	9.00	6.50
5	60	50	1.55	0.75	5.65	10.15	5.90	6.25

Table C.23: L-lower Percent coverage of 90% confidence limit for β, based on the simulations with $k = 10$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_{.}$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	6.55	6.40	6.70	13.30	7.55	6.60
1.5	15	20	7.11	7.05	6.10	13.76	7.96	7.16
2	15	20	6.75	6.05	6.95	14.30	7.00	6.95
5	15	20	6.35	6.10	6.10	16.30	6.85	7.50
1	30	20	6.35	6.25	6.90	12.70	7.20	6.35
1.5	30	20	6.30	6.80	6.05	12.80	7.15	6.35
2	30	20	6.40	6.50	7.75	13.30	7.50	6.60
5	30	20	5.45	5.30	6.30	14.25	6.90	7.00
1	60	20	6.20	5.60	6.40	14.40	8.05	6.25
1.5	60	20	7.15	5.75	5.60	13.91	8.65	7.15
2	60	20	6.60	5.90	6.00	13.75	7.60	6.95
5	60	20	5.25	4.00	5.55	13.80	8.20	7.80
1	15	35	5.76	6.10	5.85	12.21	6.71	5.76
1.5	15	35	6.95	7.05	5.35	12.35	7.20	7.10
2	15	35	6.95	6.80	5.85	11.65	6.65	7.05
5	15	35	6.75	5.45	4.75	15.41	7.40	9.60
1	30	35	6.05	6.05	5.25	13.80	7.55	6.10
1.5	30	35	5.85	5.15	5.90	12.25	7.65	6.05
2	30	35	6.65	6.55	5.40	12.30	7.10	6.90
5	30	35	4.90	3.40	5.75	12.30	7.40	7.60
1	60	35	5.95	5.25	6.65	13.45	8.95	5.95
1.5	60	35	5.05	6.85	5.90	12.00	6.50	5.05
2	60	35	5.60	5.55	6.15	12.95	7.25	5.90
5	60	35	4.45	3.20	5.60	11.60	7.60	6.70
1	15	50	5.35	4.15	5.75	12.86	6.65	5.40
1.5	15	50	6.10	6.20	6.20	11.40	6.35	6.15
2	15	50	7.45	6.60	4.30	12.51	6.80	7.80
5	15	50	5.25	4.10	5.20	14.45	5.55	7.75
1	30	50	5.55	5.10	5.30	12.70	7.70	5.65
1.5	30	50	6.65	5.05	5.35	12.80	8.00	6.70
2	30	50	6.95	5.90	5.80	12.65	8.00	7.15
5	30	50	4.75	2.85	5.10	12.55	7.50	8.45
1	60	50	5.25	5.05	5.50	12.65	6.70	5.35
1.5	60	50	5.35	5.65	5.90	10.80	6.45	5.30
2	60	50	5.45	5.35	6.10	10.75	7.00	5.75
5	60	50	3.20	3.05	5.05	9.85	6.45	5.55

Table C.24: U-Upper Percent coverage of 90% confidence limit for β, based on the simulations with $k = 1$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	6.15	7.40	5.90	12.80	10.65	6.15
1.5	15	20	5.60	5.20	6.55	14.40	11.50	6.00
2	15	20	4.90	4.30	5.70	13.35	10.75	5.20
5	15	20	6.55	12.80	5.30	13.90	9.70	5.40
1	30	20	7.10	7.10	6.45	14.25	11.65	7.05
1.5	30	20	5.15	5.05	5.70	11.75	10.35	5.35
2	30	20	5.60	5.55	5.15	13.30	12.20	5.90
5	30	20	8.40	15.05	4.60	14.25	10.55	6.00
1	60	20	6.15	8.00	6.20	12.90	10.40	6.45
1.5	60	20	5.60	6.60	6.90	12.85	11.35	5.80
2	60	20	5.20	5.45	5.50	12.45	10.85	5.25
5	60	20	7.15	16.70	5.20	11.30	8.85	5.35
1	15	35	5.40	6.15	6.10	11.60	11.05	5.65
1.5	15	35	4.65	4.85	6.60	13.10	10.65	4.70
2	15	35	4.85	4.85	5.10	13.85	11.90	4.90
5	15	35	6.75	16.10	4.65	14.65	8.95	4.20
1	30	35	5.55	5.00	5.80	12.25	11.20	5.30
1.5	30	35	5.20	4.75	5.75	13.00	11.65	5.60
2	30	35	5.40	4.85	4.30	14.25	11.05	5.75
5	30	35	7.80	20.05	4.20	12.95	9.20	4.85
1	60	35	5.95	6.35	5.85	12.20	11.95	6.10
1.5	60	35	6.05	5.10	5.25	13.20	11.20	6.15
2	60	35	5.40	5.95	4.95	12.65	11.15	5.60
5	60	35	8.45	24.70	4.50	11.80	8.65	5.05
1	15	50	5.30	5.90	6.10	11.95	11.20	5.80
1.5	15	50	4.90	5.25	3.95	12.95	11.70	5.30
2	15	50	4.95	5.25	4.75	14.25	11.45	5.55
5	15	50	6.70	19.60	4.75	13.85	8.40	3.35
1	30	50	5.55	6.05	5.95	12.00	10.85	5.25
1.5	30	50	5.80	3.85	5.15	12.65	11.55	5.90
2	30	50	5.75	5.35	4.45	13.60	12.00	6.35
5	30	50	8.25	27.25	5.00	13.20	8.55	4.60
1	60	50	5.20	5.25	5.10	11.20	10.10	5.85
1.5	60	50	5.15	4.05	5.30	12.10	11.80	5.80
2	60	50	5.65	5.30	4.85	11.65	10.85	5.35
5	60	50	9.70	32.25	5.05	11.80	9.35	4.60

Table C.25: U-Upper Percent coverage of 90% confidence limit for β, based on the simulations with $k = 3$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	5.20	6.35	5.95	11.20	6.65	5.15
1.5	15	20	4.75	5.85	5.55	13.70	8.80	4.65
2	15	20	5.25	6.00	6.55	14.50	8.35	5.10
5	15	20	7.90	10.75	6.10	16.95	8.55	4.65
1	30	20	6.30	6.00	6.25	13.65	9.15	6.50
1.5	30	20	5.85	4.90	6.20	13.50	9.40	6.05
2	30	20	5.75	6.25	6.70	13.75	8.10	5.75
5	30	20	8.95	13.05	5.45	15.15	7.80	5.15
1	60	20	5.80	6.70	5.90	13.20	9.70	5.80
1.5	60	20	5.45	5.80	7.10	11.10	7.75	5.75
2	60	20	6.35	5.80	6.25	14.10	8.60	6.25
5	60	20	10.00	14.30	4.50	14.05	8.00	5.75
1	15	35	6.10	5.05	5.05	12.85	9.35	6.30
1.5	15	35	6.15	4.35	5.75	14.95	9.75	6.05
2	15	35	4.30	5.80	5.95	13.20	7.35	4.05
5	15	35	8.35	11.45	5.65	15.35	7.50	4.05
1	30	35	5.75	5.95	5.40	13.00	8.70	5.95
1.5	30	35	5.65	4.90	5.35	14.05	8.60	5.70
2	30	35	4.60	5.85	5.20	12.50	8.10	4.50
5	30	35	11.50	15.50	5.35	16.35	8.95	4.85
1	60	35	6.45	5.10	5.25	12.75	9.45	6.25
1.5	60	35	5.55	6.05	6.00	12.50	9.60	5.45
2	60	35	4.90	5.85	5.55	12.40	8.00	4.90
5	60	35	10.00	20.55	5.60	11.80	6.50	3.95
1	15	50	5.55	6.00	5.55	11.70	7.75	5.45
1.5	15	50	4.85	4.80	6.00	13.90	9.00	4.75
2	15	50	4.35	4.90	6.35	14.40	8.15	4.30
5	15	50	8.80	14.00	5.80	16.90	7.35	3.75
1	30	50	5.05	4.95	5.85	11.55	8.30	4.90
1.5	30	50	4.95	4.00	4.85	13.70	9.50	4.70
2	30	50	5.00	5.45	4.40	14.90	8.35	4.55
5	30	50	10.45	19.90	5.80	14.40	7.25	3.95
1	60	50	4.80	5.50	4.65	11.15	8.60	4.85
1.5	60	50	5.85	4.70	5.05	13.70	8.80	6.20
2	60	50	5.60	6.05	5.10	13.20	8.20	5.20
5	60	50	11.90	23.40	5.10	12.25	6.60	4.70

Table C.26: U-Upper Percent coverage of 90% confidence limit for β, based on the simulations with $k = 10$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	6.95	6.10	6.25	13.40	7.60	7.05
1.5	15	20	5.71	5.35	6.15	14.01	7.51	5.96
2	15	20	6.00	5.95	6.40	14.65	7.65	5.90
5	15	20	6.30	7.15	6.35	15.55	6.85	4.80
1	30	20	5.95	7.60	5.35	12.75	7.50	5.95
1.5	30	20	6.00	5.90	6.70	14.55	7.90	6.05
2	30	20	5.50	5.50	5.55	14.85	7.00	5.45
5	30	20	5.90	8.35	6.30	15.20	6.05	4.35
1	60	20	6.70	5.55	6.50	13.90	7.60	6.75
1.5	60	20	5.35	6.40	6.25	14.01	7.60	5.20
2	60	20	6.00	6.75	6.50	14.20	7.80	5.80
5	60	20	7.45	7.95	6.90	13.75	7.40	5.10
1	15	35	5.21	5.80	5.35	11.91	6.01	5.21
1.5	15	35	5.20	5.85	6.45	13.20	7.05	5.25
2	15	35	4.65	4.75	5.25	13.60	6.60	4.55
5	15	35	4.80	6.70	5.85	16.06	6.30	3.55
1	30	35	5.95	5.70	5.90	13.10	7.25	5.90
1.5	30	35	5.15	5.15	5.40	14.35	7.60	5.10
2	30	35	4.90	5.60	5.80	13.25	6.95	4.70
5	30	35	8.35	8.75	5.20	15.95	7.35	5.35
1	60	35	5.85	5.60	5.05	12.05	7.80	5.80
1.5	60	35	4.50	5.50	5.55	12.30	6.65	4.45
2	60	35	5.10	6.45	5.80	13.20	7.95	5.05
5	60	35	7.95	9.85	5.65	14.25	7.00	4.85
1	15	50	6.20	5.35	5.85	12.56	6.80	6.25
1.5	15	50	5.30	5.00	5.70	14.80	7.15	5.35
2	15	50	4.25	4.50	5.05	13.26	6.50	4.15
5	15	50	5.80	7.50	6.00	18.15	6.95	3.40
1	30	50	5.45	6.35	4.85	12.10	7.45	5.35
1.5	30	50	4.55	3.95	6.20	12.85	6.90	4.60
2	30	50	4.85	5.70	5.45	13.75	7.15	4.50
5	30	50	6.65	8.40	6.10	15.15	6.15	4.10
1	60	50	4.95	5.05	5.15	11.90	7.25	5.00
1.5	60	50	4.95	5.50	5.15	12.90	6.90	5.00
2	60	50	5.25	5.20	5.75	13.75	8.05	4.95
5	60	50	10.00	11.10	6.95	16.15	8.20	4.75

Table C.27: T-Total Percent coverage of 90% confidence limits for β, based on the simulations with $k = 1$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	12.65	15.80	13.20	24.45	21.30	12.85
1.5	15	20	12.35	13.30	12.65	26.15	20.75	12.40
2	15	20	11.30	12.20	13.10	23.75	20.80	11.65
5	15	20	11.05	14.70	12.30	23.40	15.65	11.85
1	30	20	12.85	14.00	13.60	25.75	21.65	12.60
1.5	30	20	11.85	11.65	12.00	24.55	21.05	11.45
2	30	20	12.95	11.55	11.90	25.65	21.35	13.20
5	30	20	12.80	17.40	11.45	24.20	16.50	11.75
1	60	20	12.30	15.20	12.65	24.90	22.05	12.35
1.5	60	20	11.80	13.60	14.95	24.10	21.15	11.60
2	60	20	12.25	12.20	13.60	25.20	20.80	12.30
5	60	20	11.25	17.60	13.00	23.05	14.60	11.55
1	15	35	10.95	13.50	11.70	22.65	21.60	11.25
1.5	15	35	10.35	11.35	11.95	23.75	19.90	10.35
2	15	35	12.55	11.95	11.15	25.30	22.90	12.75
5	15	35	11.75	17.00	10.40	25.55	15.85	12.25
1	30	35	11.20	11.50	12.10	24.70	21.85	11.00
1.5	30	35	11.70	11.60	12.65	24.05	22.35	11.50
2	30	35	11.10	12.05	11.20	24.70	20.95	11.30
5	30	35	11.35	20.85	10.70	24.25	15.55	11.65
1	60	35	12.75	11.25	12.85	24.95	24.15	12.70
1.5	60	35	11.75	11.20	11.30	24.10	21.15	12.00
2	60	35	11.55	12.50	11.50	24.05	21.20	11.55
5	60	35	12.00	25.55	11.10	23.70	15.35	11.65
1	15	50	9.75	12.90	12.65	23.20	22.35	10.50
1.5	15	50	11.70	12.20	9.55	23.85	21.60	11.90
2	15	50	11.65	10.90	10.65	24.25	21.10	12.25
5	15	50	11.85	20.25	11.00	24.30	14.85	12.20
1	30	50	10.45	11.45	11.65	22.85	21.40	10.05
1.5	30	50	11.60	10.60	10.65	22.80	21.70	11.55
2	30	50	12.70	11.70	10.00	25.40	21.20	12.95
5	30	50	11.85	27.85	11.05	22.90	15.00	10.65
1	60	50	10.50	10.60	10.55	22.70	22.65	11.25
1.5	60	50	11.30	11.40	11.55	22.20	22.60	11.70
2	60	50	11.00	11.35	10.35	22.30	21.90	10.85
5	60	50	13.00	32.50	11.30	22.85	15.75	11.05

Table C.28: T-Total Percent coverage of 90% confidence limits for β, based on the simulations with $k = 3$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	μ	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	11.65	12.85	11.70	25.05	15.10	11.65
1.5	15	20	11.65	11.95	11.45	25.70	16.60	11.25
2	15	20	12.50	12.75	12.10	27.60	16.55	12.40
5	15	20	13.20	13.90	12.40	29.15	16.20	13.10
1	30	20	13.30	12.75	12.50	27.50	18.10	13.45
1.5	30	20	13.40	11.95	12.20	26.65	18.80	13.55
2	30	20	12.85	13.85	12.30	26.65	16.15	13.35
5	30	20	13.75	15.95	11.75	28.45	15.40	13.10
1	60	20	12.30	13.15	12.30	26.80	18.70	12.20
1.5	60	20	11.75	12.85	13.10	23.55	16.70	11.95
2	60	20	13.15	11.25	12.65	26.60	16.40	13.05
5	60	20	13.95	16.50	10.25	25.80	14.85	12.95
1	15	35	12.50	10.15	11.55	26.40	18.15	12.80
1.5	15	35	12.55	10.95	12.45	26.85	18.30	12.50
2	15	35	11.65	12.15	11.30	24.40	15.15	11.35
5	15	35	12.45	13.80	11.55	25.90	14.35	11.75
1	30	35	11.60	11.00	10.60	24.45	16.80	11.65
1.5	30	35	12.20	10.90	11.15	26.65	17.05	12.40
2	30	35	10.40	9.75	10.95	23.70	15.65	10.55
5	30	35	15.05	17.45	10.50	28.25	16.00	12.15
1	60	35	12.15	10.25	11.70	24.80	18.20	12.25
1.5	60	35	12.10	11.90	10.95	25.15	18.45	11.70
2	60	35	10.20	11.40	11.00	23.40	15.85	10.55
5	60	35	12.60	22.05	11.15	23.50	13.80	10.25
1	15	50	10.65	11.50	10.75	24.55	16.40	11.00
1.5	15	50	11.55	11.70	11.15	25.50	16.85	11.50
2	15	50	11.80	11.40	11.70	25.10	16.25	11.95
5	15	50	11.95	15.70	11.10	26.20	13.40	11.45
1	30	50	10.90	10.45	10.75	22.95	16.60	10.85
1.5	30	50	11.20	9.75	10.85	24.90	17.75	11.10
2	30	50	11.40	9.70	9.65	26.15	16.30	11.35
5	30	50	13.05	21.20	10.70	24.90	15.10	11.00
1	60	50	10.25	11.05	10.55	23.85	17.10	10.40
1.5	60	50	11.60	10.35	10.05	24.25	16.50	11.80
2	60	50	11.60	11.25	10.70	26.55	17.20	11.70
5	60	50	13.45	24.15	10.75	22.40	12.50	10.95

Table C.29: T-Total Percent coverage of 90% confidence limits for β, based on the simulations with $k = 10$. Other simulation factors are listed in columns 1-3. $\rho = \exp(\beta)$.

ρ	$\mu_.$	Nobs	Cond NB 1	Cond NB 2	Cond NB 3	Conc NB	Bin GLM_OD	Bin GLMM
1	15	20	13.50	12.51	12.95	26.70	15.15	13.65
1.5	15	20	12.81	12.41	12.25	27.78	15.47	13.11
2	15	20	12.75	12.01	13.35	28.95	14.65	12.85
5	15	20	12.65	13.26	12.46	31.85	13.70	12.30
1	30	20	12.30	13.85	12.25	25.45	14.70	12.30
1.5	30	20	12.30	12.70	12.75	27.35	15.05	12.40
2	30	20	11.90	12.00	13.30	28.15	14.50	12.05
5	30	20	11.35	13.65	12.60	29.45	12.95	11.35
1	60	20	12.90	11.15	12.90	28.30	15.65	13.00
1.5	60	20	12.51	12.15	11.85	27.91	16.26	12.36
2	60	20	12.60	12.65	12.50	27.95	15.40	12.75
5	60	20	12.70	11.95	12.45	27.55	15.60	12.90
1	15	35	10.96	11.90	11.20	24.12	12.71	10.96
1.5	15	35	12.15	12.91	11.80	25.55	14.25	12.35
2	15	35	11.60	11.55	11.11	25.25	13.25	11.60
5	15	35	11.56	12.15	10.60	31.47	13.71	13.16
1	30	35	12.00	11.75	11.15	26.90	14.80	12.00
1.5	30	35	11.00	10.30	11.30	26.60	15.25	11.15
2	30	35	11.55	12.15	11.20	25.55	14.05	11.60
5	30	35	13.25	12.15	10.95	28.25	14.75	12.95
1	60	35	11.80	10.85	11.70	25.50	16.75	11.75
1.5	60	35	9.55	12.35	11.45	24.30	13.15	9.50
2	60	35	10.70	12.00	11.95	26.15	15.20	10.95
5	60	35	12.40	13.05	11.25	25.85	14.60	11.55
1	15	50	11.56	9.50	11.60	25.41	13.46	11.66
1.5	15	50	11.40	11.20	11.90	26.20	13.50	11.50
2	15	50	11.71	11.10	9.35	25.76	13.31	11.96
5	15	50	11.05	11.60	11.20	32.60	12.50	11.15
1	30	50	11.00	11.45	10.15	24.80	15.15	11.00
1.5	30	50	11.20	9.00	11.55	25.65	14.90	11.30
2	30	50	11.80	11.60	11.25	26.40	15.15	11.65
5	30	50	11.40	11.25	11.20	27.70	13.65	12.55
1	60	50	10.20	10.10	10.65	24.55	13.95	10.35
1.5	60	50	10.30	11.15	11.05	23.70	13.35	10.30
2	60	50	10.70	10.55	11.85	24.50	15.05	10.70
5	60	50	13.20	14.15	12.00	26.00	14.65	10.30

Appendix D

Figures

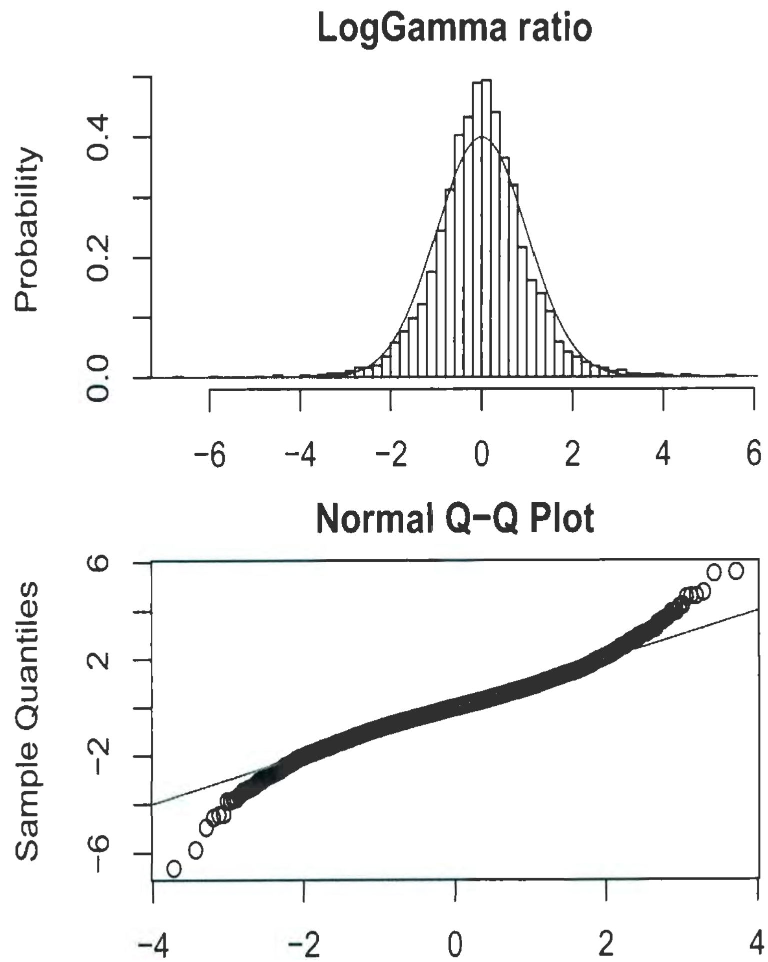

Figure D.1: Top: log gamma ratio divided by the standard deviation. Bottom: qq plot. The straight line is for the standard normal distribution.

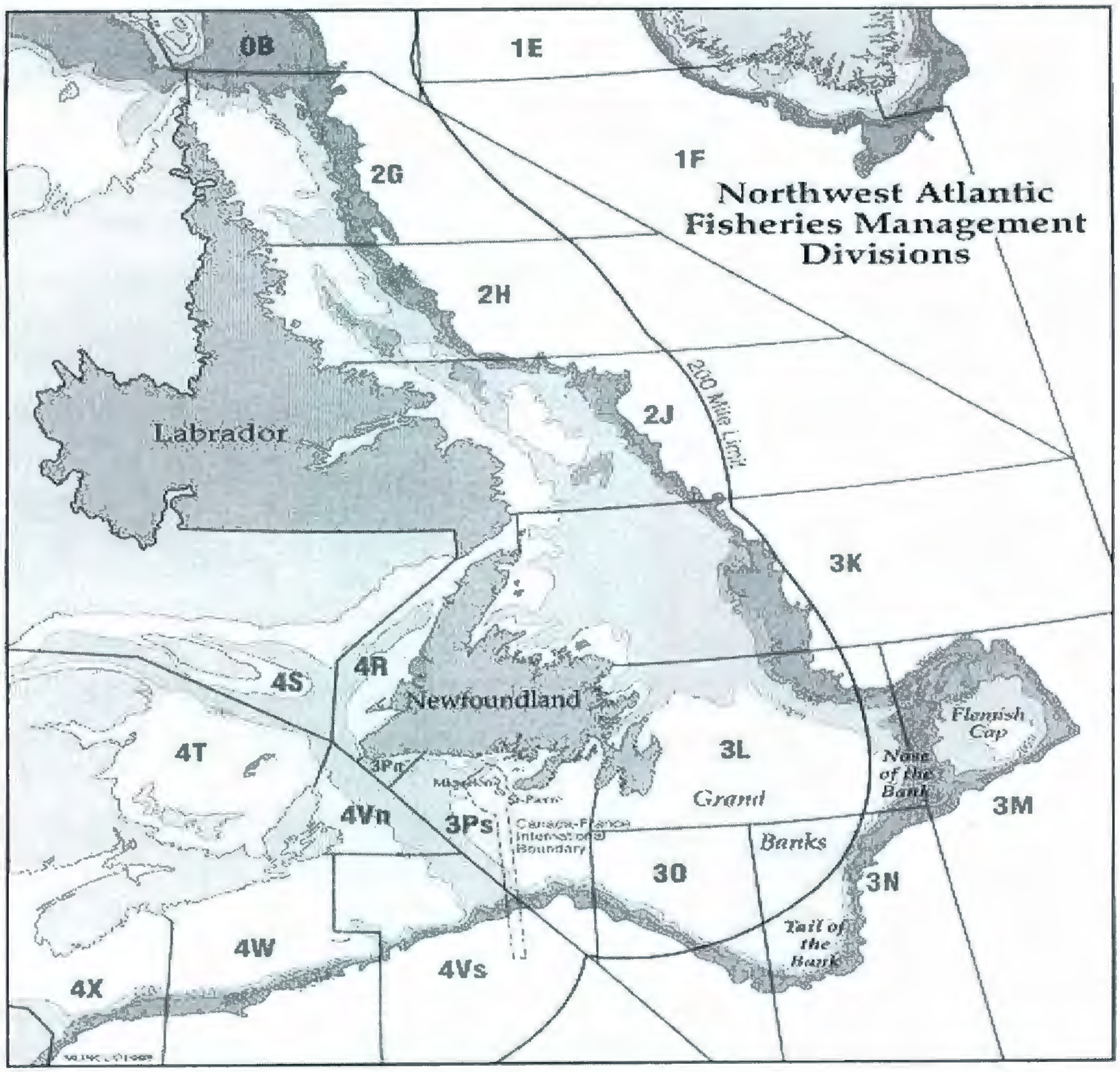

Figure D.2: Northwest Atlantic Fisheries Organization (NAFO) management Divisions.

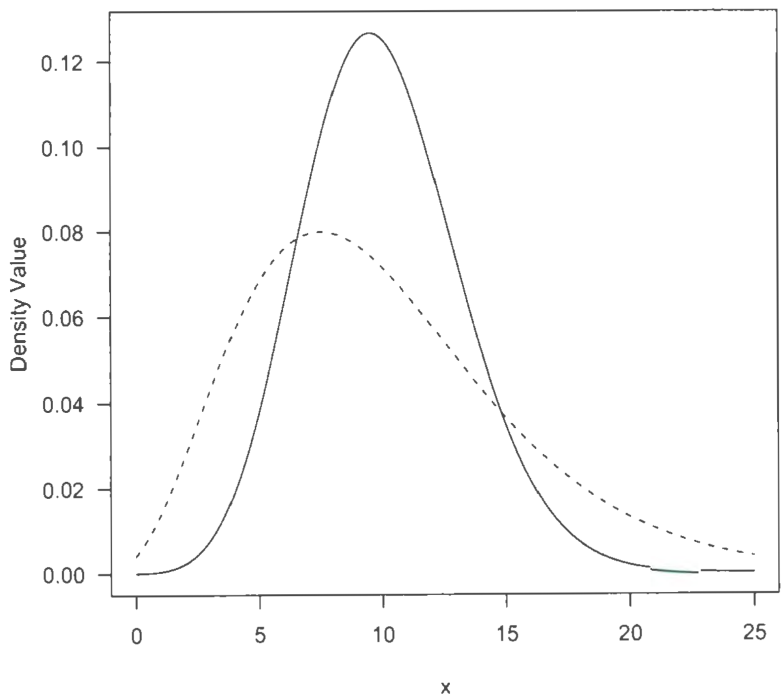

Figure D.3: NB compared with Poisson distibution. Negative Binomial density function is presented by dotted curve. Poisson density function is presented by solid curve.

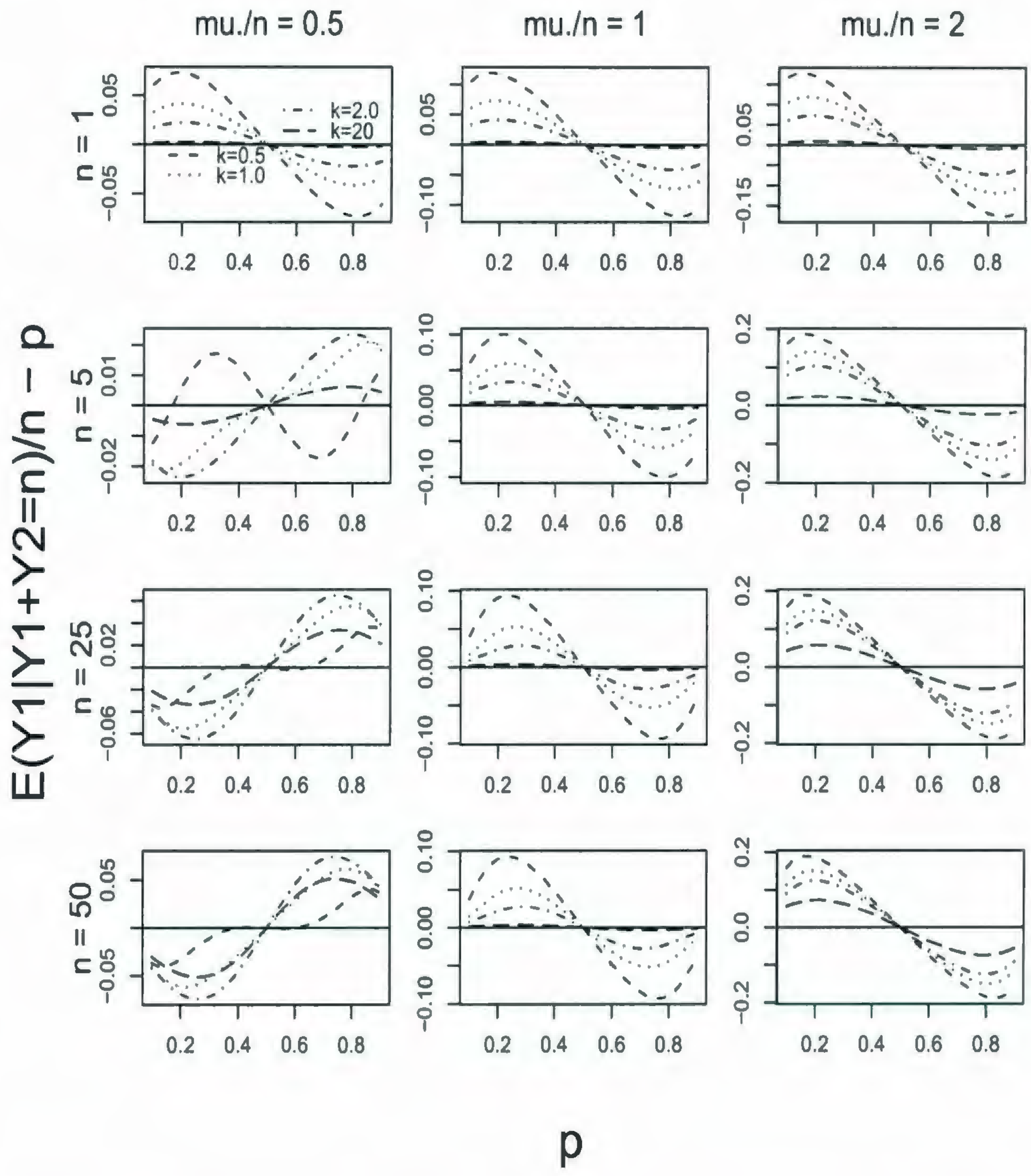

Figure D.4: Difference in the expected NB fraction $E(Y_1|Y_.)/n$ and the Binomial probability p, as a function of $p = \mu_1/\mu_{..}$. Each panel shows results for a value of n and $\mu_.$, which is shown as a ratio of n. Different line types are for different values of the NB k, which are shown in the top left-hand panel. The y-scale may differ between panels.

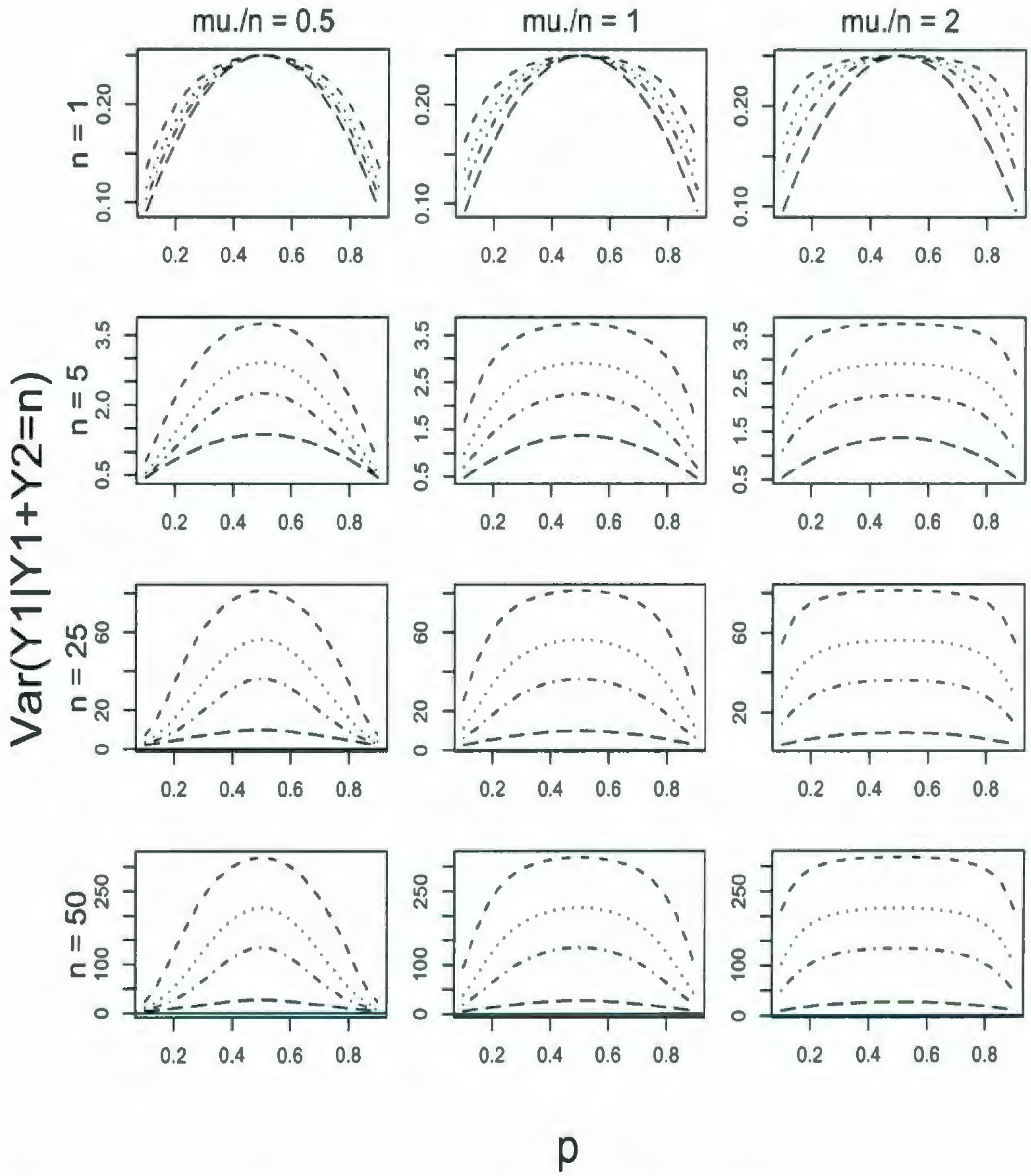

Figure D.5: Variance of the conditional NB distribution. See Figure D.4 for details.

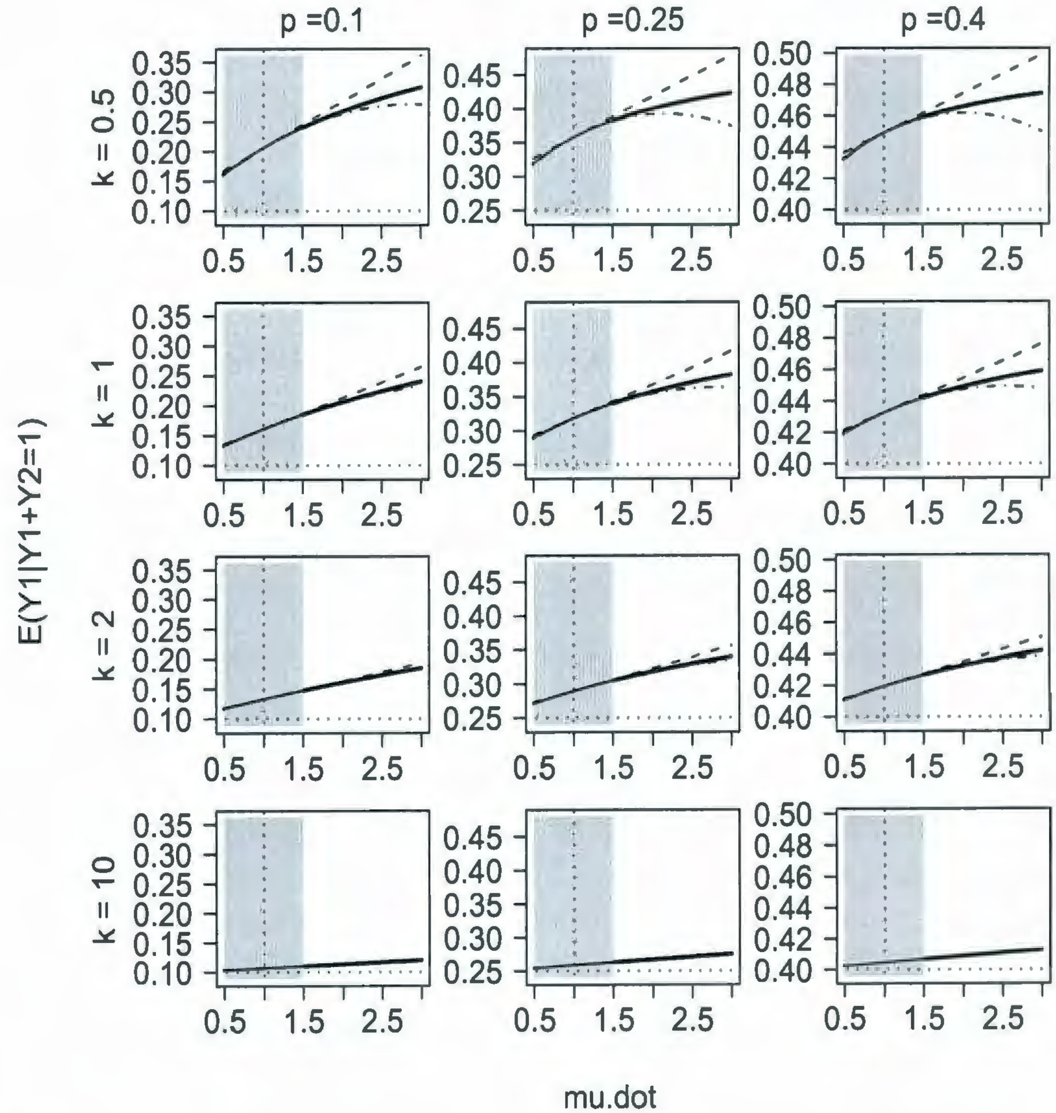

Figure D.6: The heavy solid line is the NB conditional expectation versus of $\mu_{.}$ when n = 1. The dashed and dotted-dashed curves are Taylor's series approximations. Each panel shows results for a choice of $p = \mu_1/\mu.$ and k. The horizontal dotted line denotes the Binomial expectation. A vertical dotted line at n is shown, and the shaded region covers $\pm 50\%$ of n.

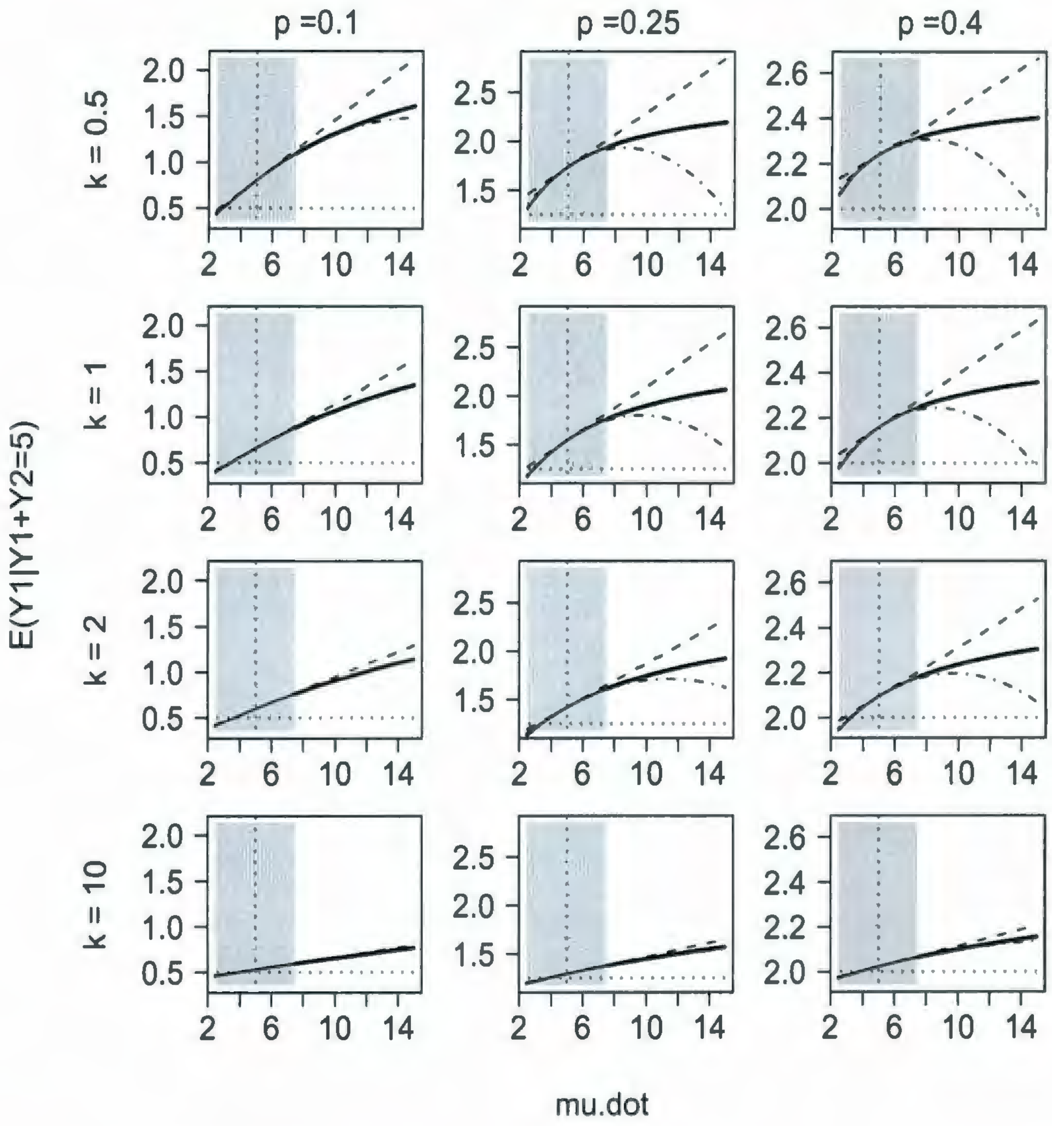

Figure D.7: NB conditional expectation when n = 5. See Figure D.6 for details.

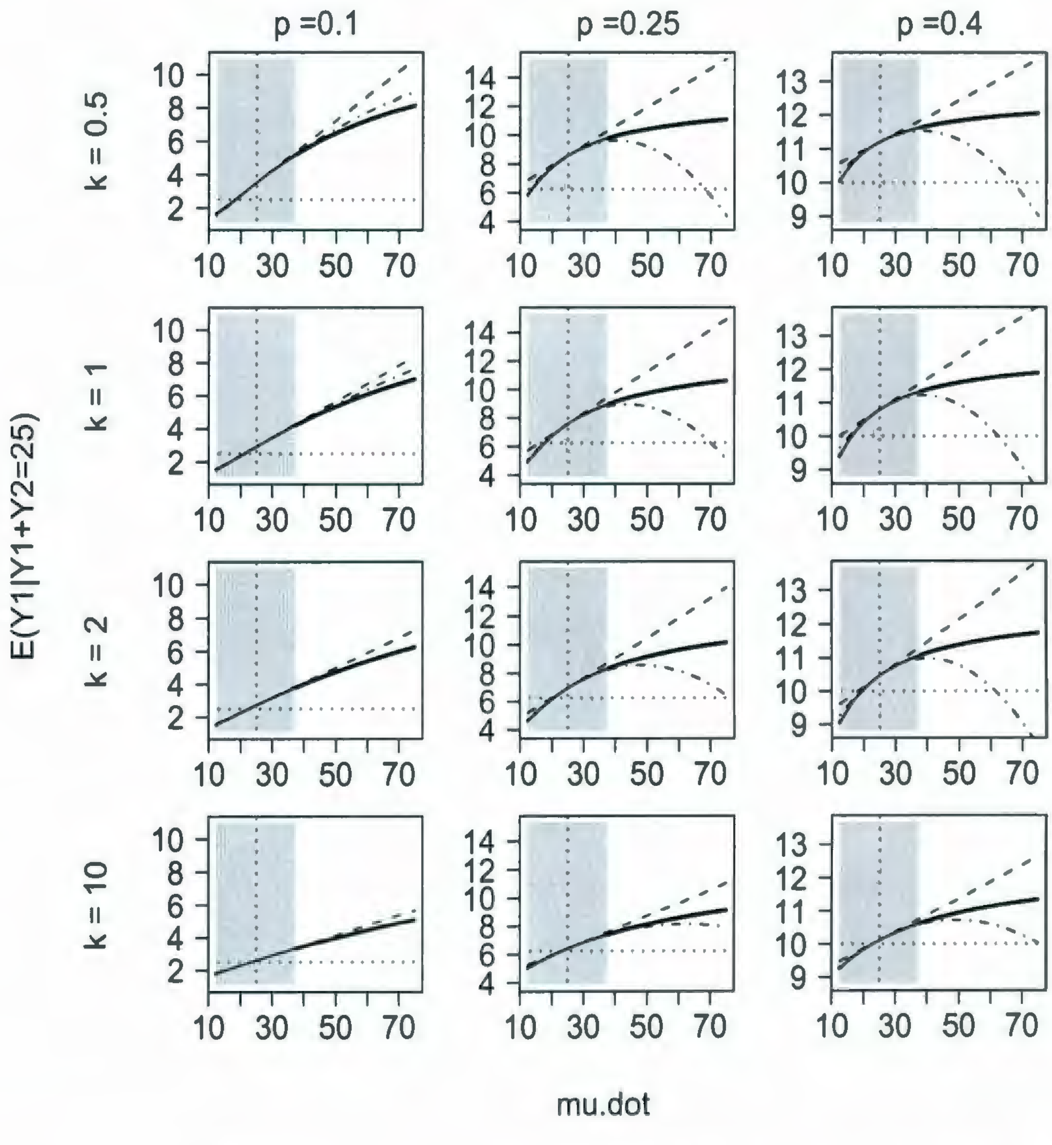

Figure D.8: NB conditional expectation when n = 25. See Figure D.6 for details.

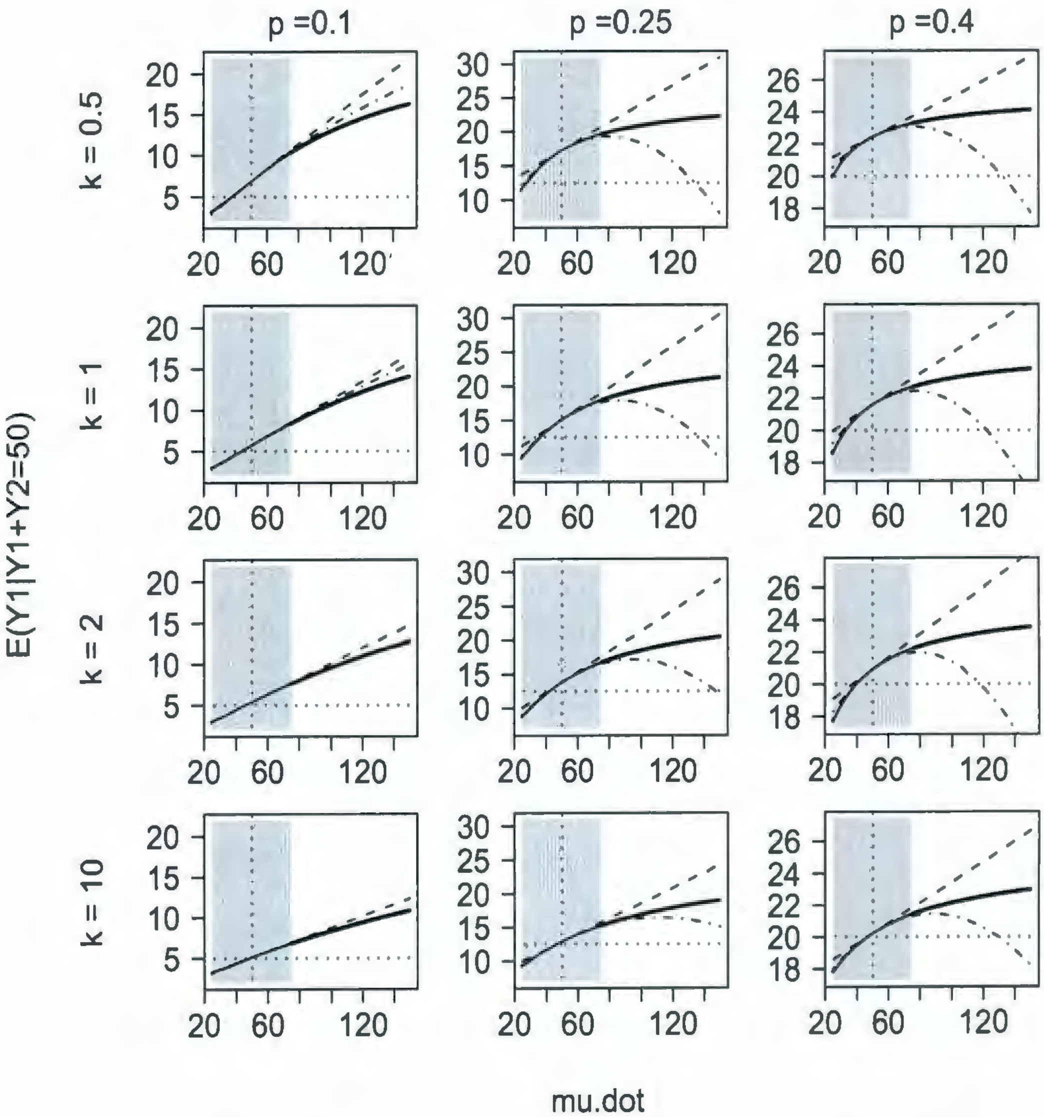

Figure D.9: NB conditional expectation when n = 50. See Figure D.6 for details.

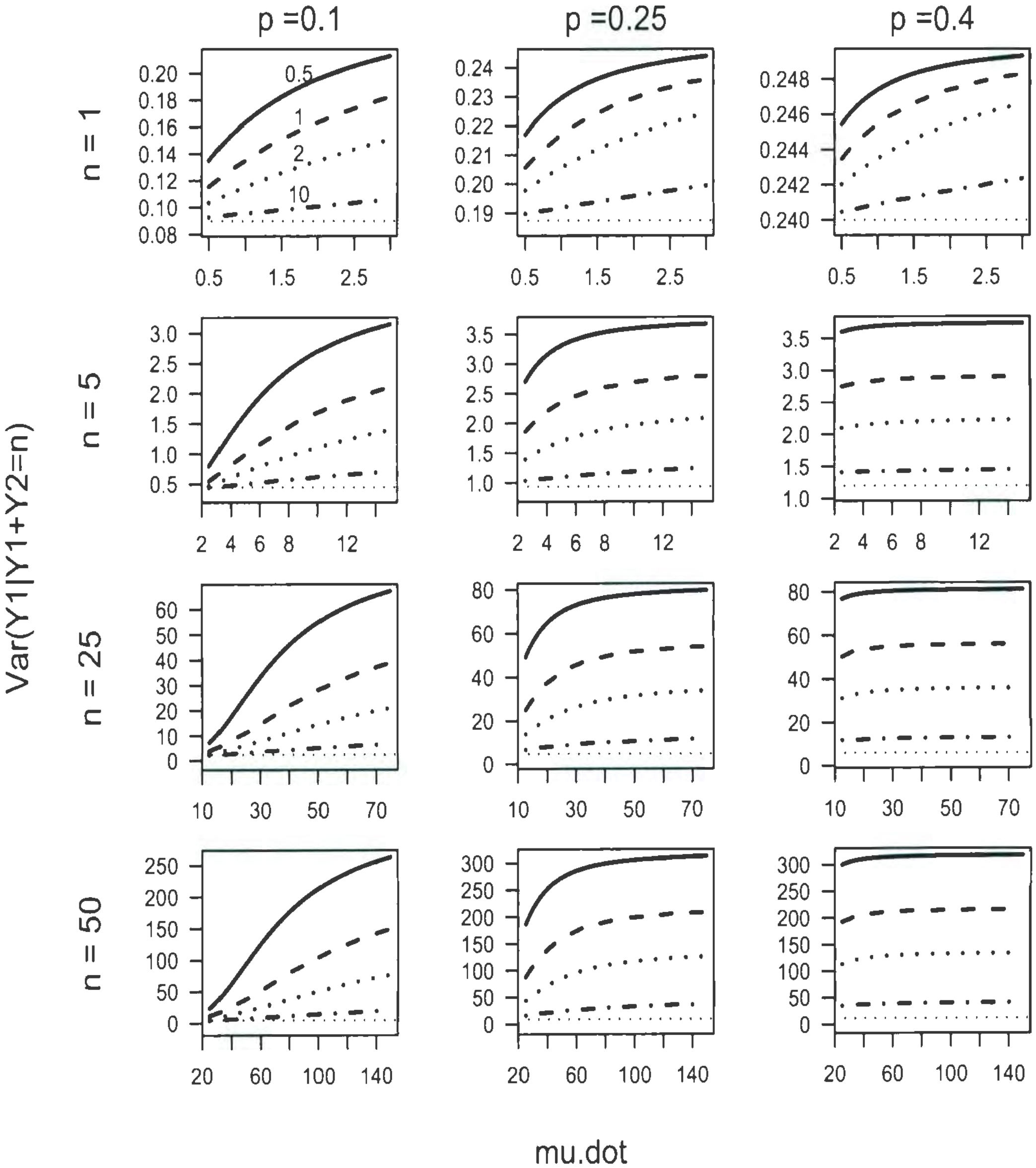

Figure D.10: Conditional NB variance versus $\mu_.$. Each panel shows the results for a choice of n and $p = \mu_1/\mu_.$. Line types correspond to different values of the NB k parameter, which are shown in the top left-hand panel. The Binomial variance is shown as the horizontal dotted line.

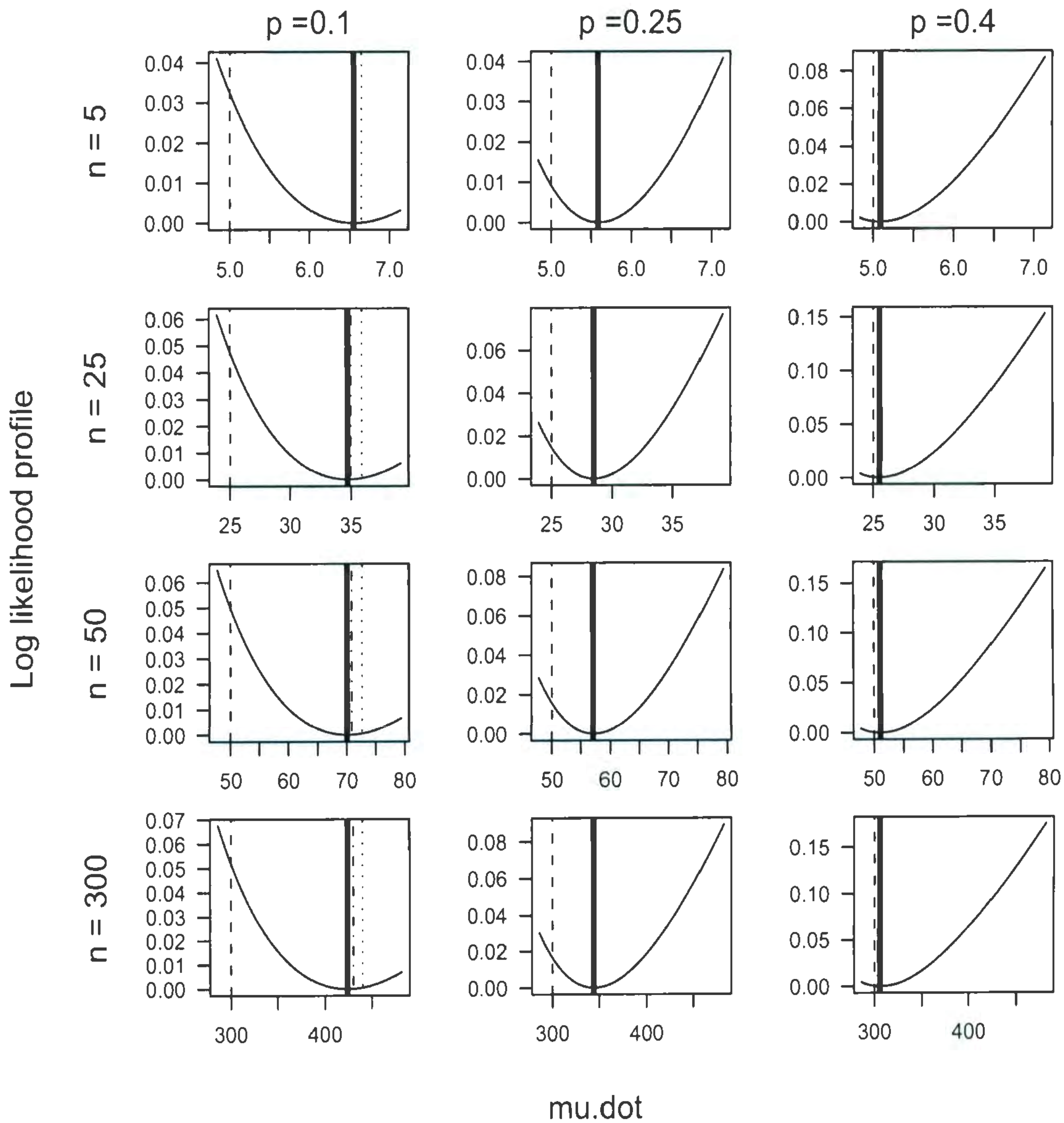

Figure D.11: Loglikelihood Profile (thin solid line) of $\mu_.$ for k=0.5 and $Y_. = n$. Each panel shows the results for a choice of n and $p = \mu_1/\mu_.$. The heavy solid line denotes the mle. The dashed line represents n, while the dotted and dashed-dotted lines represent the first and second order approximations to the mle..

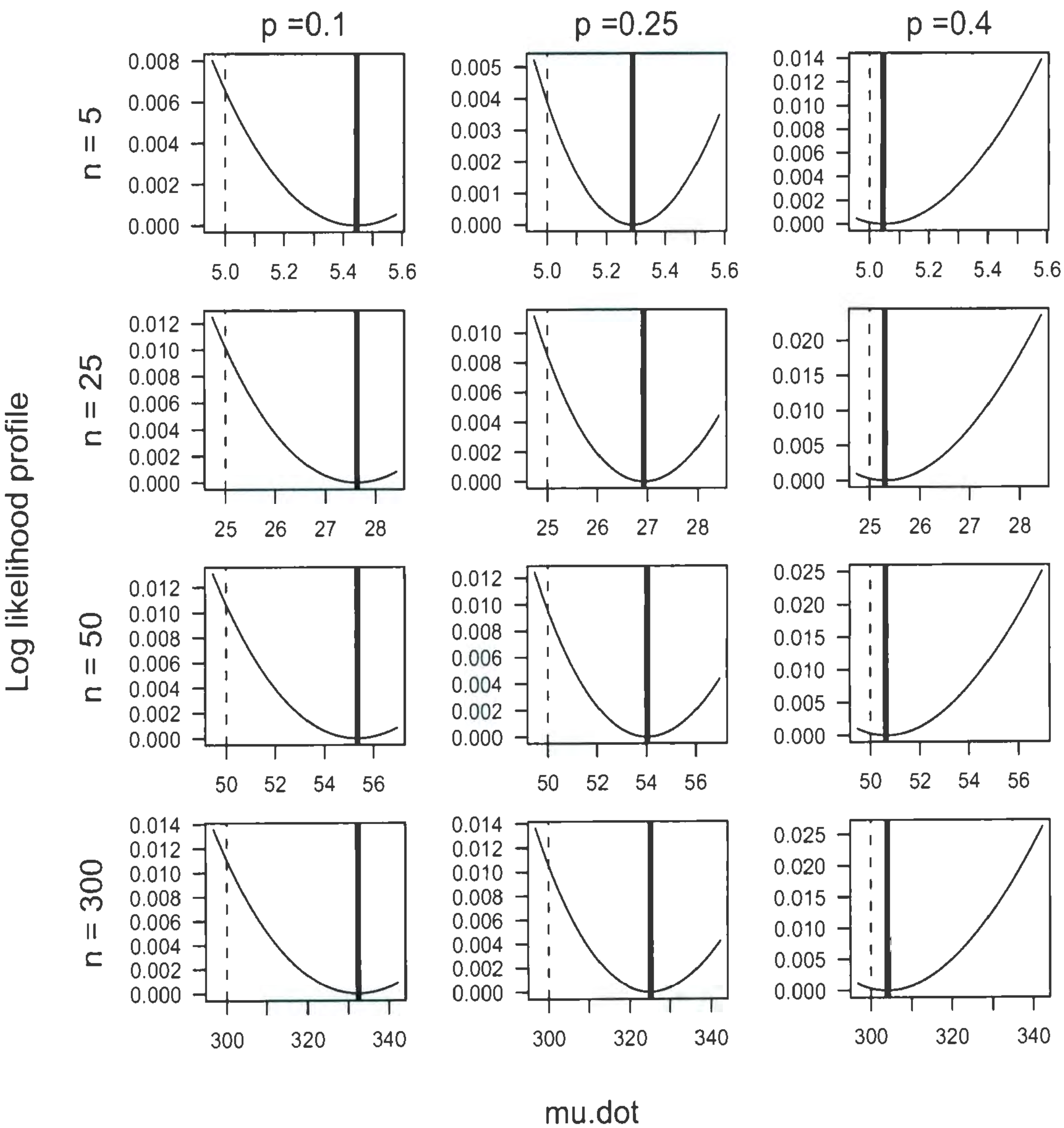

Figure D.12: Loglikelihood Profile of $\mu_{.}$ for k=1.0. See Figure D.11 for details.

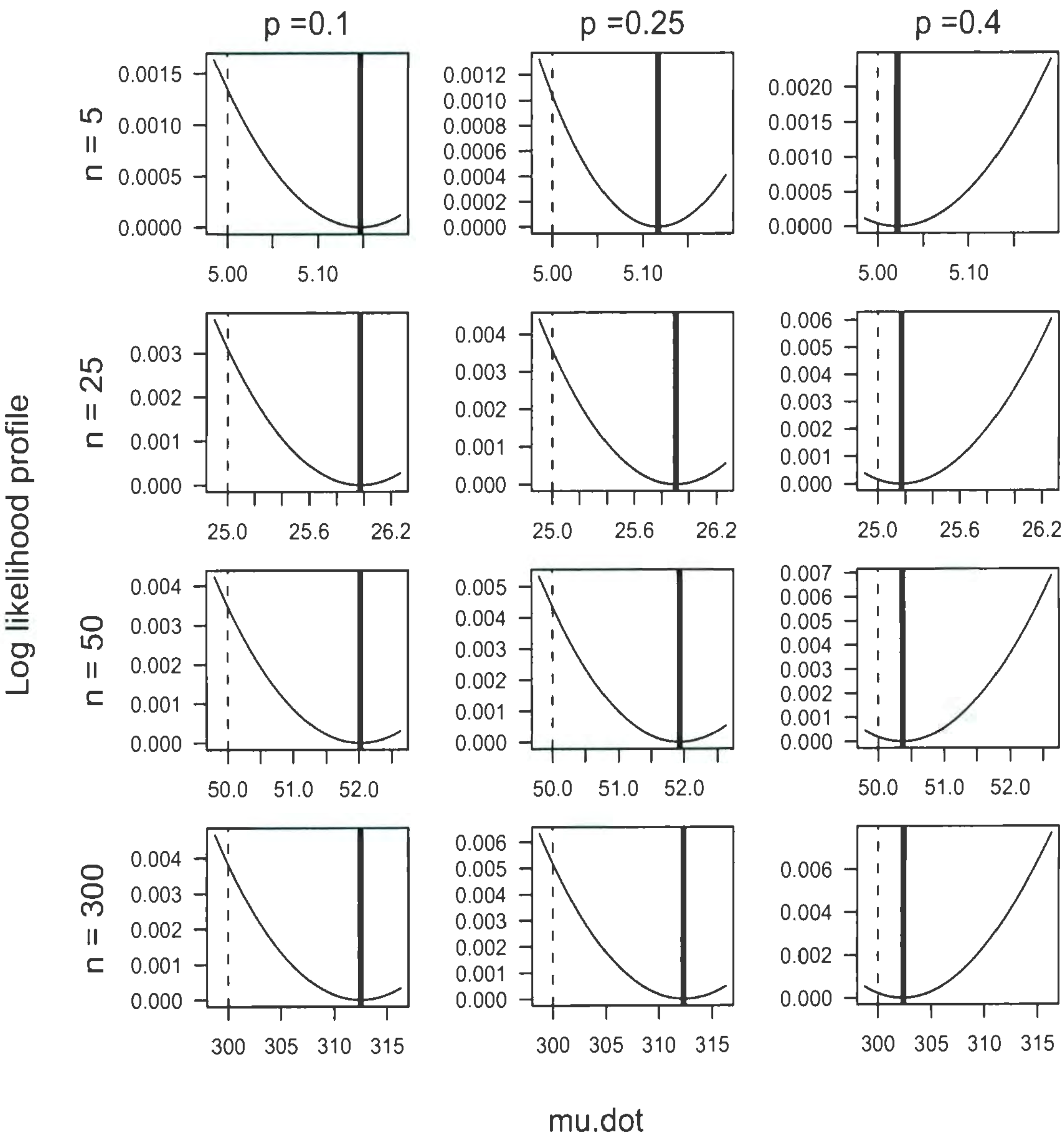

Figure D.13: Loglikelihood Profile of $\mu_\cdot$ for k=2.0. See Figure D.1 for details.

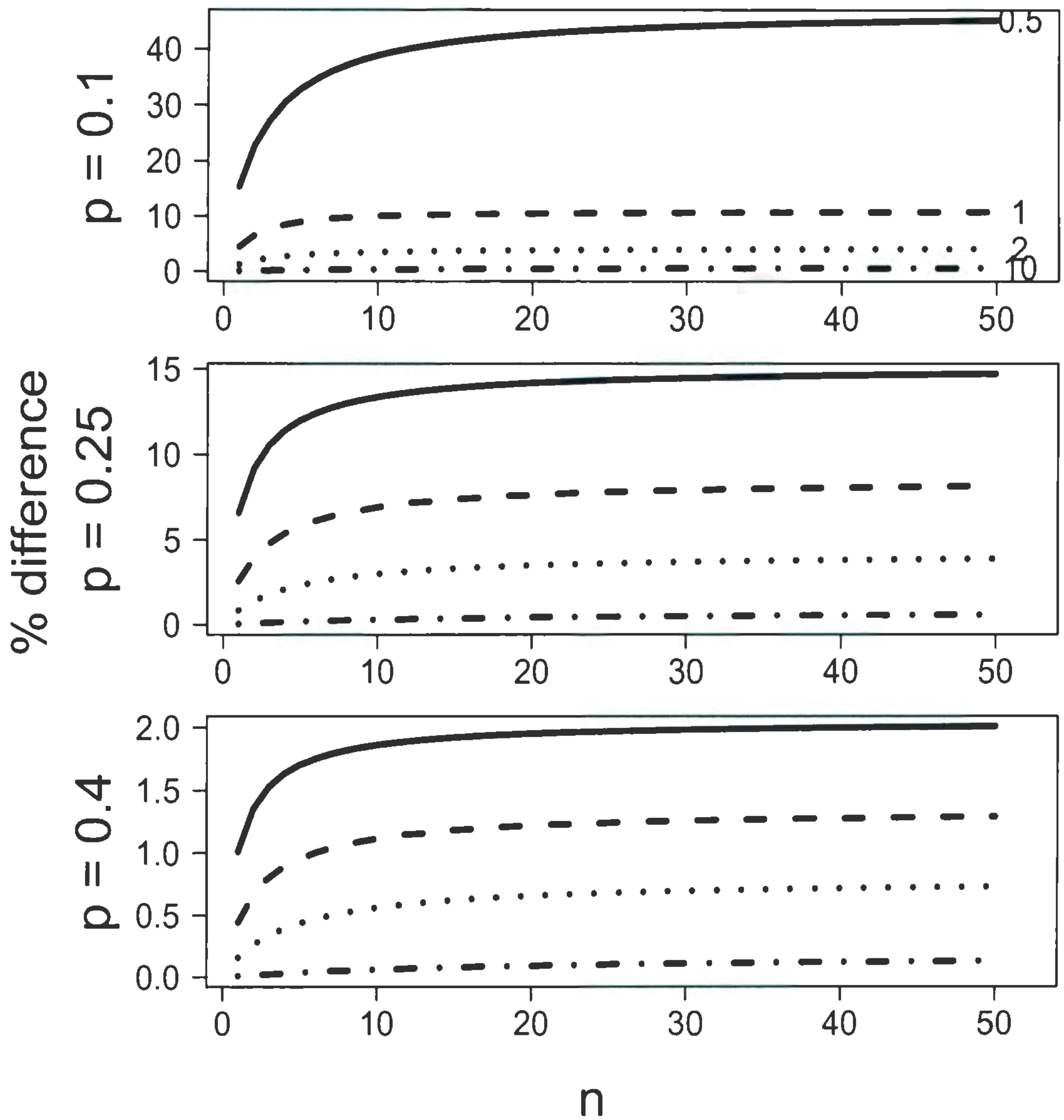

Figure D.14: Percent differences between the first-order mle of $\mu_.$ and n, versus n. Each panel shows the results for different values of $p = \mu_1/\mu_.$. Line types correspond to different values of the NB k parameter, which are shown at the right-hand side in the top panel.

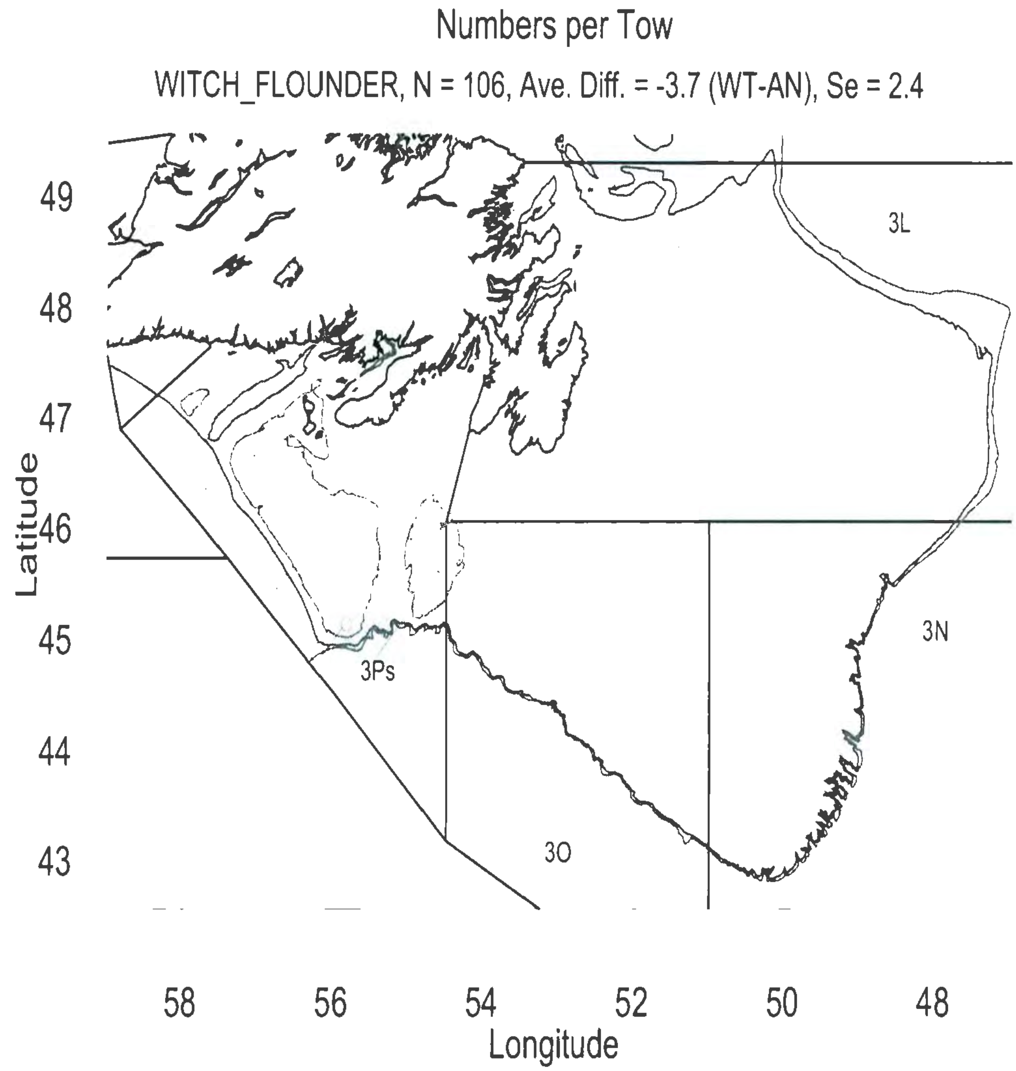

Figure D.15: Differences between witch flounder catches (WT minus AN), divided by their Poisson standard error. Catches are scaled for subsampling, and standardized for tow distance. Plotting symbols are proportional to the difference: · - negative, + - positive, ◦ - zero, o - no catch in either pair. Black ×, + are potential outliers. The number of paired sets (N), the average catch difference (Ave. Diff.), and the standard error of the differences (Se) are listed at the top. 100 m (blue), 300 m (purple), and 500 m (green) depth contours are shown.

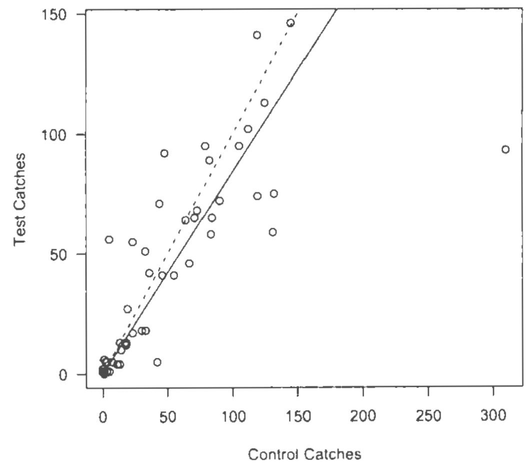

Figure D.16: The dotted line has slope of one. The solid line has a slope equal to the estimated relative efficiency (ρ) from Conditional Poisson model.

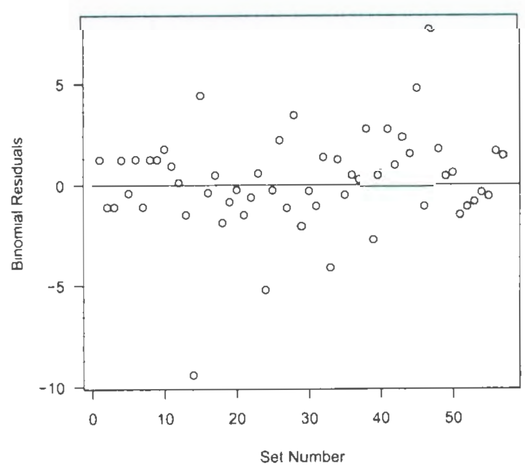

Figure D.17: Residuals for the Conditional Poisson model.

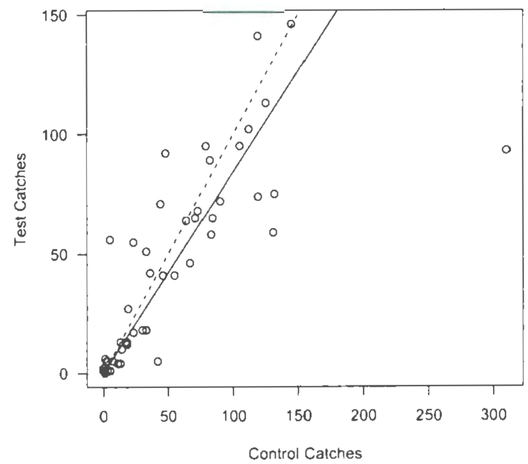

Figure D.18: The dotted line has slope of one. The solid line has a slope equal to the estimated relative efficiency (ρ) from Overdispersed Binomial model.

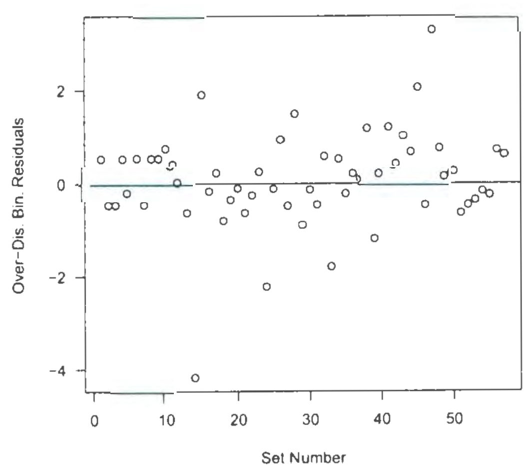

Figure D.19: Residuals for the Overdispersed Binomial model.

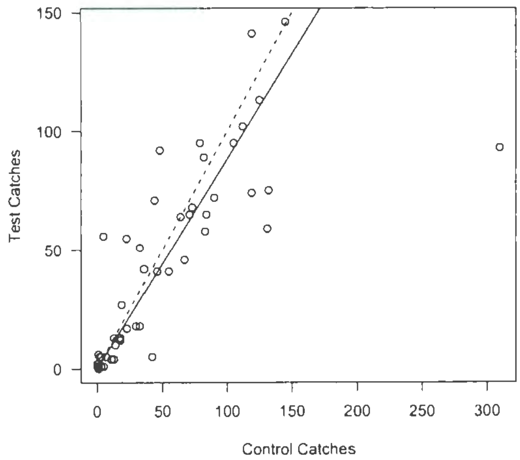

Figure D.20: The dotted line has slope of one. The solid line has a slope equal to the estimated relative efficiency (ρ) from Mixed Binomial model.

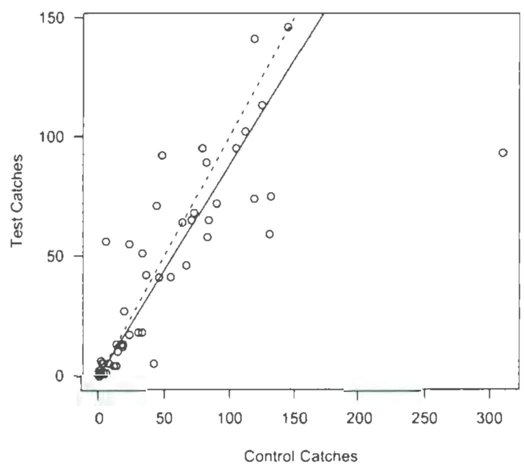

Figure D.21: The dotted line has slope of one. The solid line has a slope equal to the estimated relative efficiency (ρ) when $\mu_{\cdot} = n$ from Conditional Negative Binomial model.

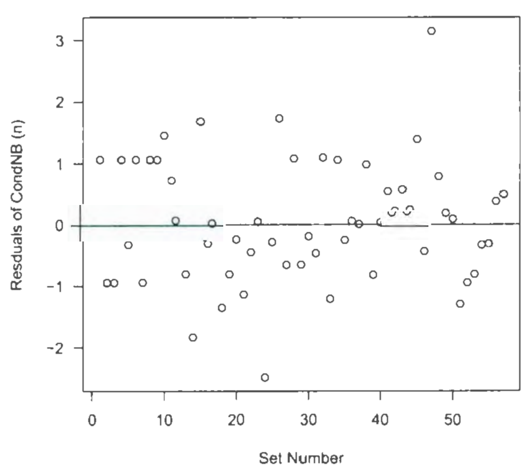

Figure D.22: Residuals for the Conditional Negative Binomial model when $\mu_{\cdot} = n$.

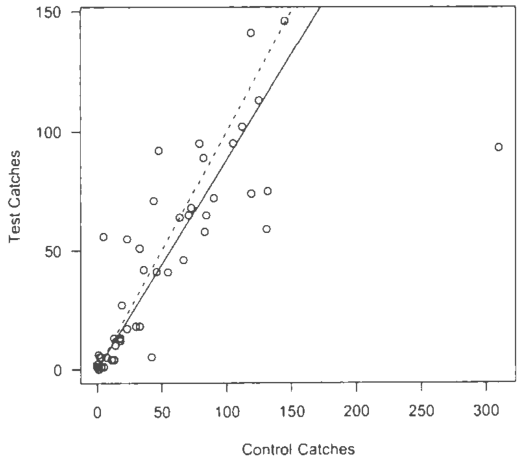

Figure D.23: The dotted line has slope of one. The solid line has a slope equal to the estimated relative efficiency (ρ) when $\mu_.$ is estimated from Conditional Negative Binomial model.

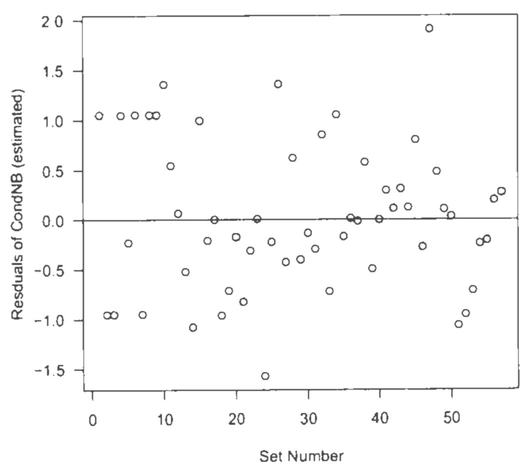

Figure D.24: Residuals for the Conditional Negative Binomial model when $\mu_.$ is estimated.

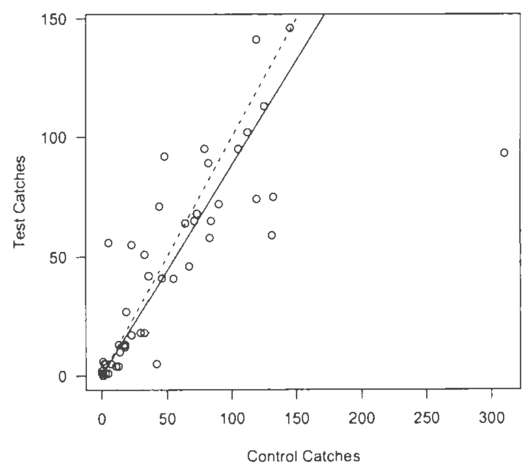

Figure D.25: The dotted line has slope of one. The solid line has a slope equal to the estimated relative efficiency (ρ) from Concentrated Negative Binomial model.

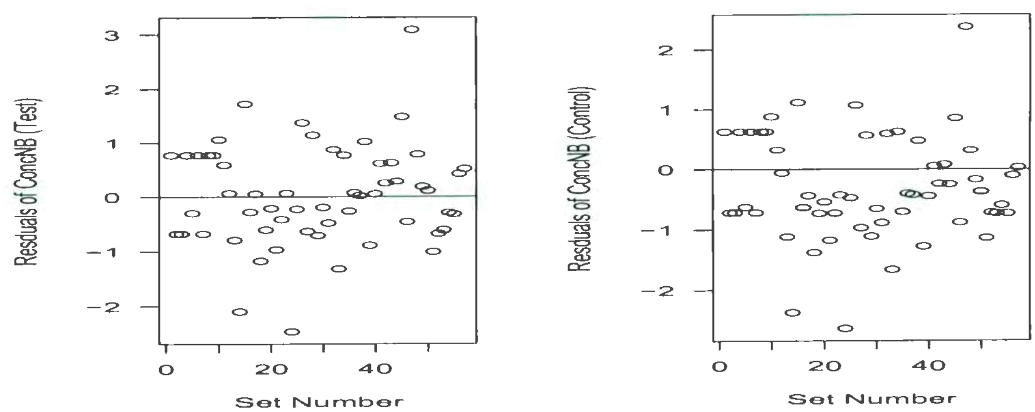

Figure D.26: Residuals for the Concentrated Negative Binomial model.

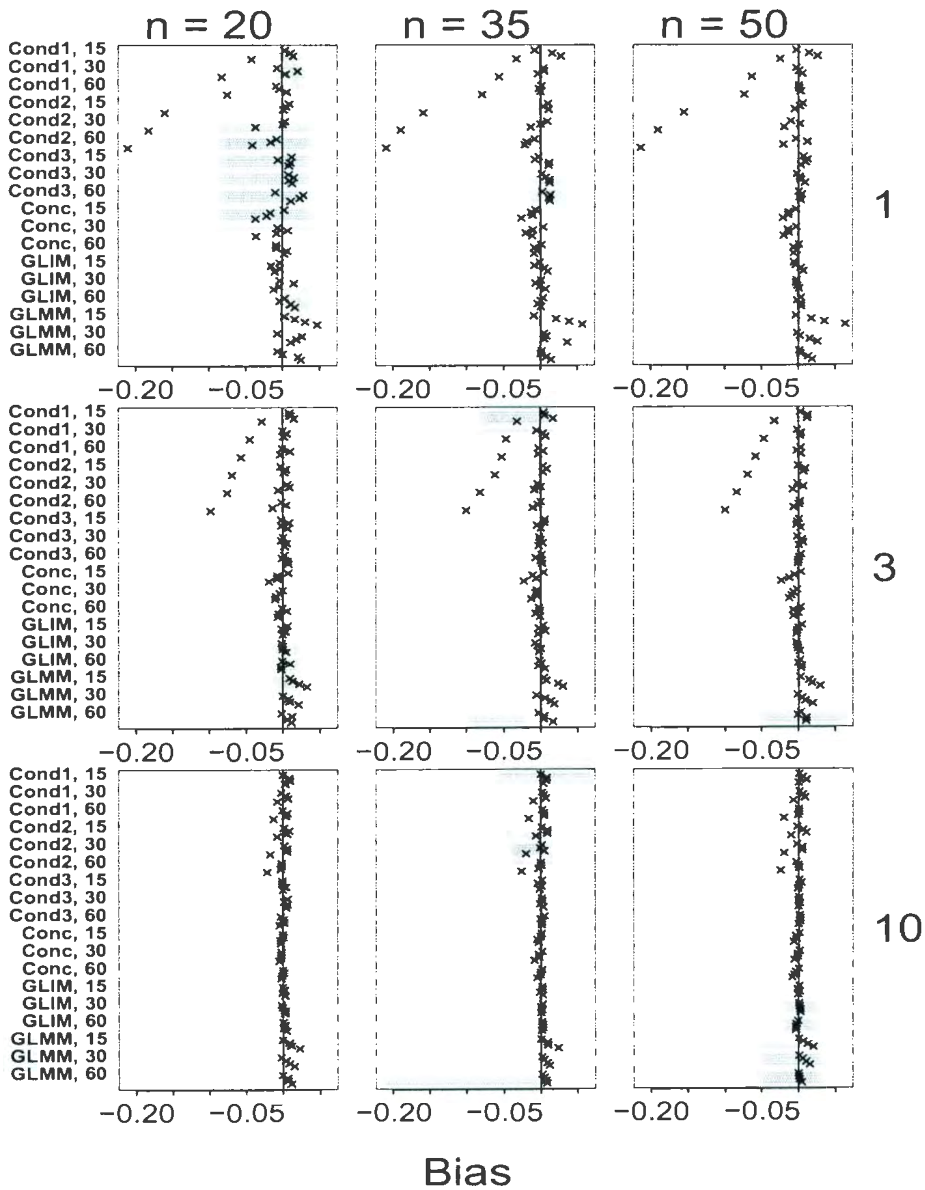

Figure D.27: Simulated bias. Each panel shows the results for a combination of n and k, which are shown at the top and right-hand side, respectively. The estimation method and value for μ are shown at the left-hand side. Each group of points corresponds to a value of $\rho = exp(\beta) = 1$ (top line), 1.5, 2, and 5 (bottom line). Conditional NB is abbreviated as Cond, and 1 indicates μ is estimated, 2 indicates $\mu = n$, and 3 indicates μ is fixed at the true value. Concentrated NB is abbreviated as Conc. Zero is shown as a solid vertical line.

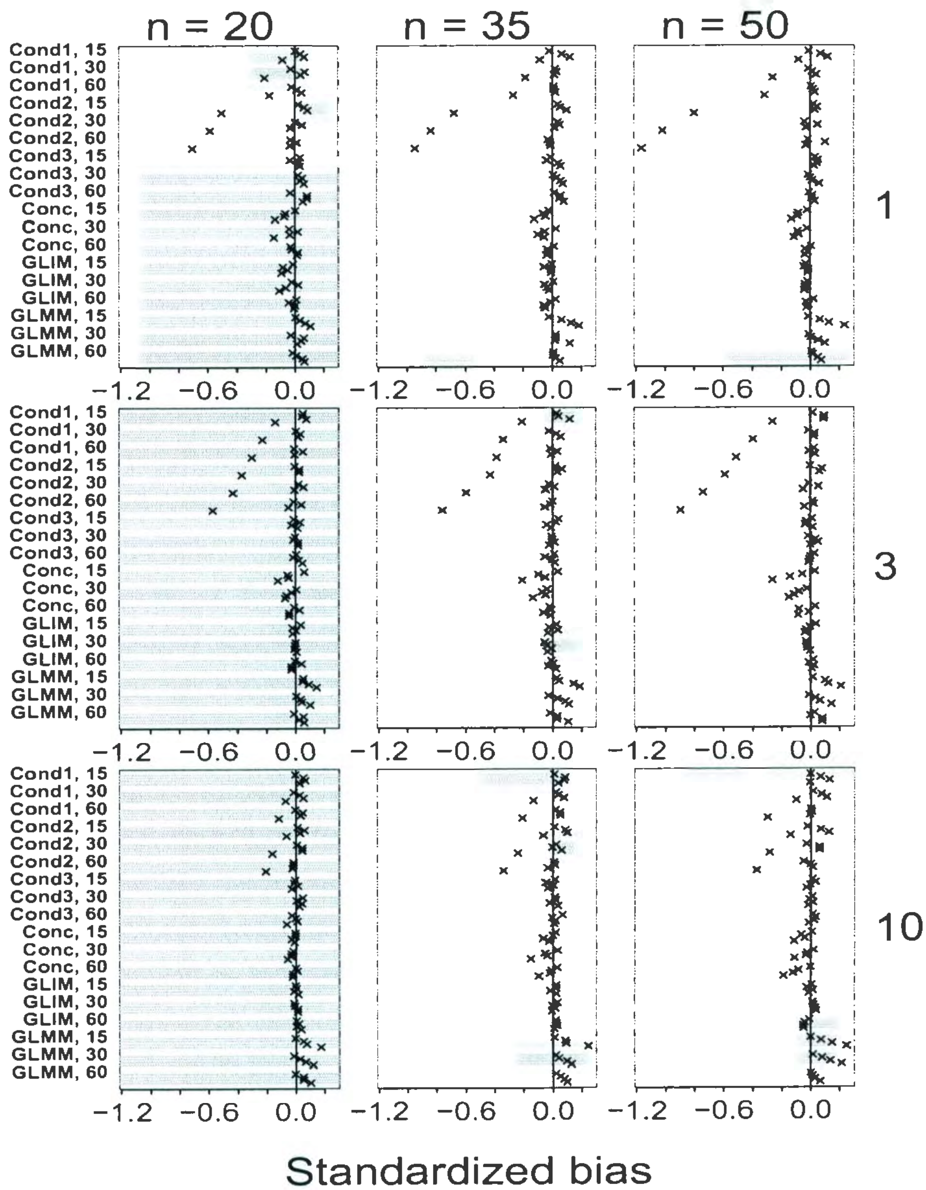

Figure D.28: Simulated standardized (e.g. by standard error) bias. See Figure D.? for details.

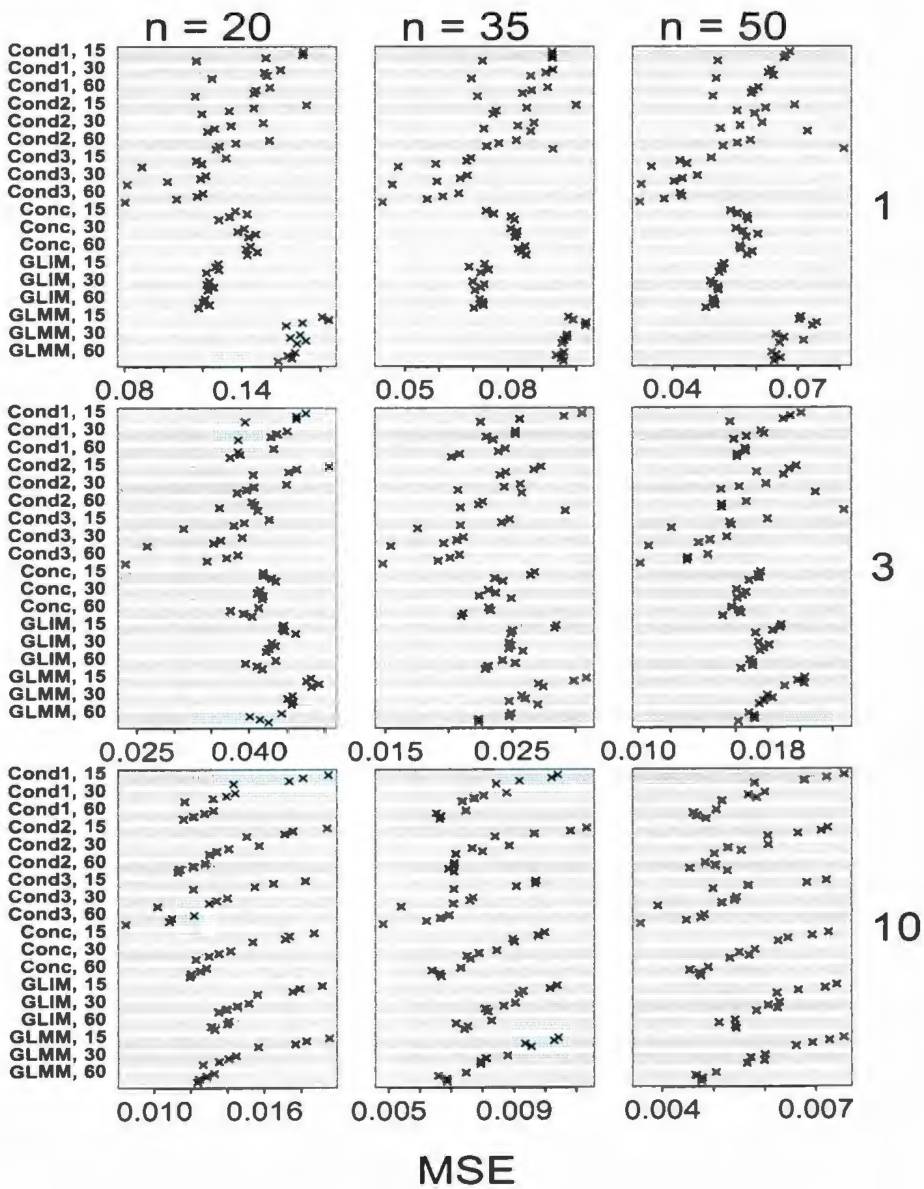

Figure D.29: Simulated MSE. See Figure D.27 for details.

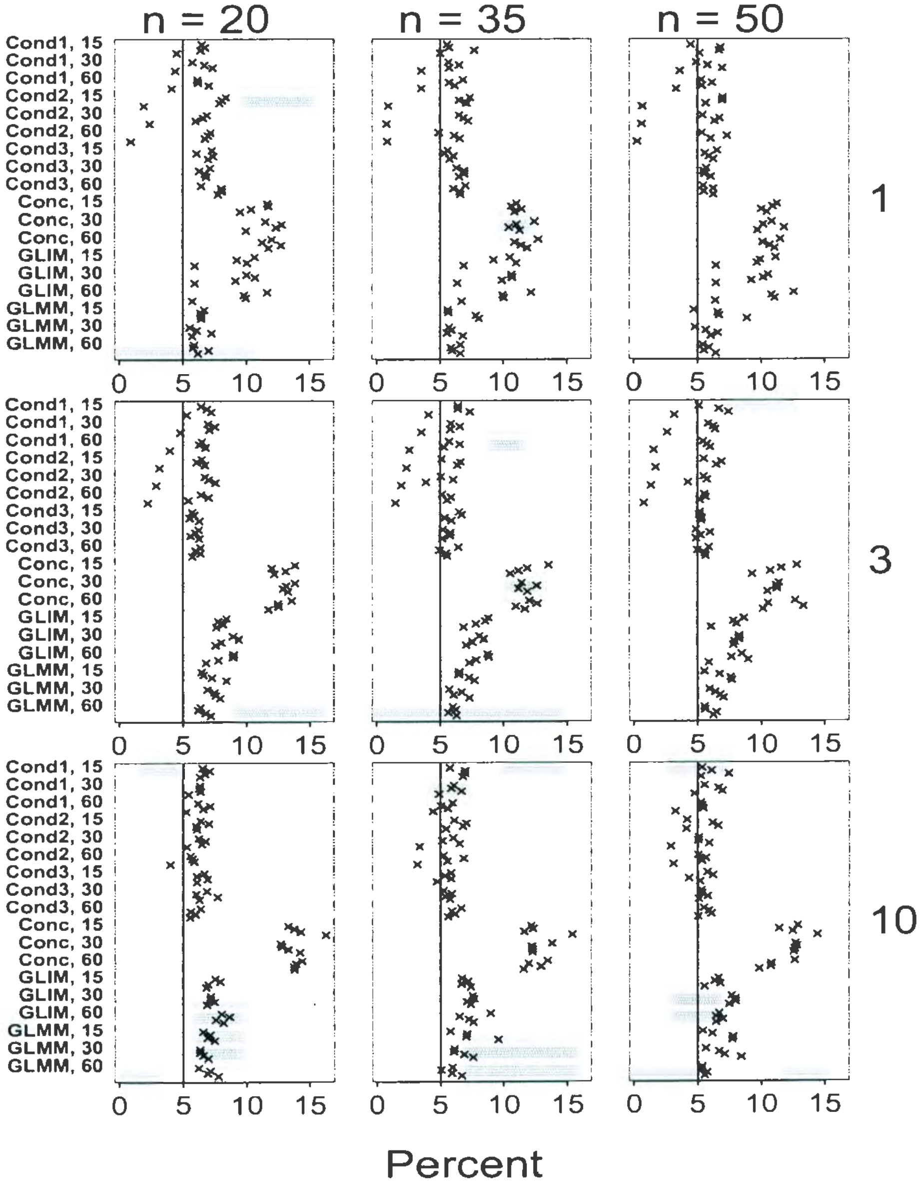

Figure D.30: Simulated lower exceedance of 90% confidence intervals. See Figure D.?? for details. 5% is shown as a solid vertical line.

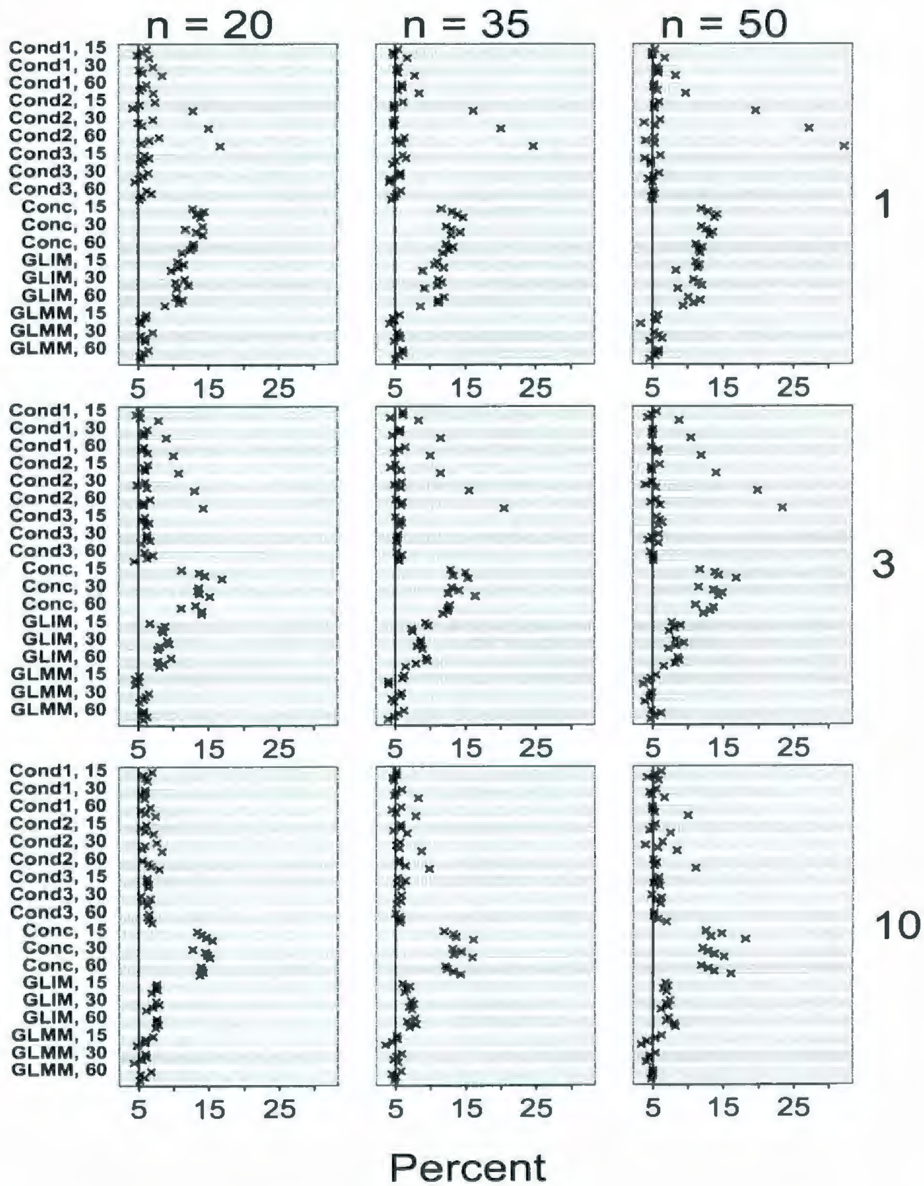

Figure D.31: Simulated upper exceedance of 90% confidence intervals. See Figure D.27 for details. 5% is shown as a solid vertical line.

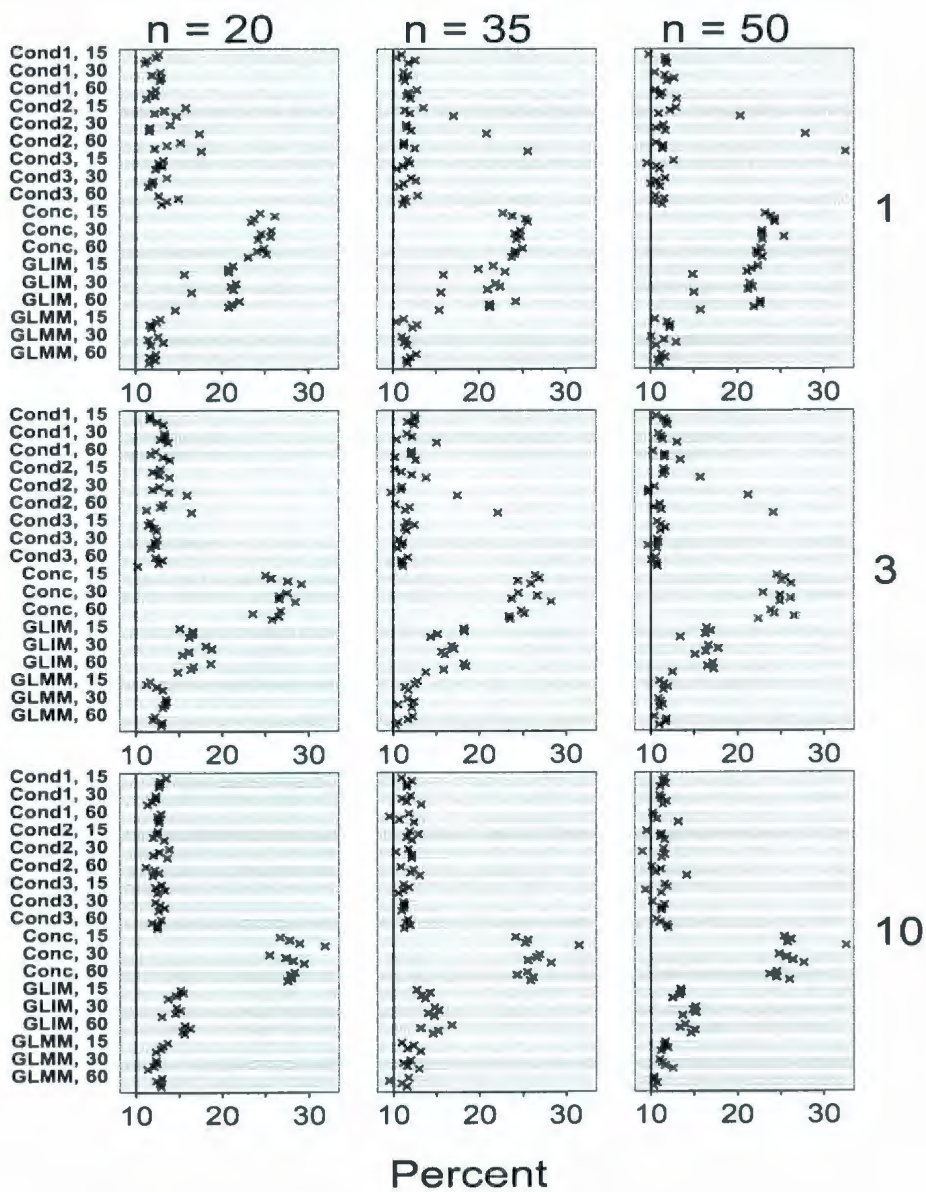

Figure D.32: Simulated total exceedance of 90% confidence intervals. See Figure D.27 for details. 10% is shown as a solid vertical line.

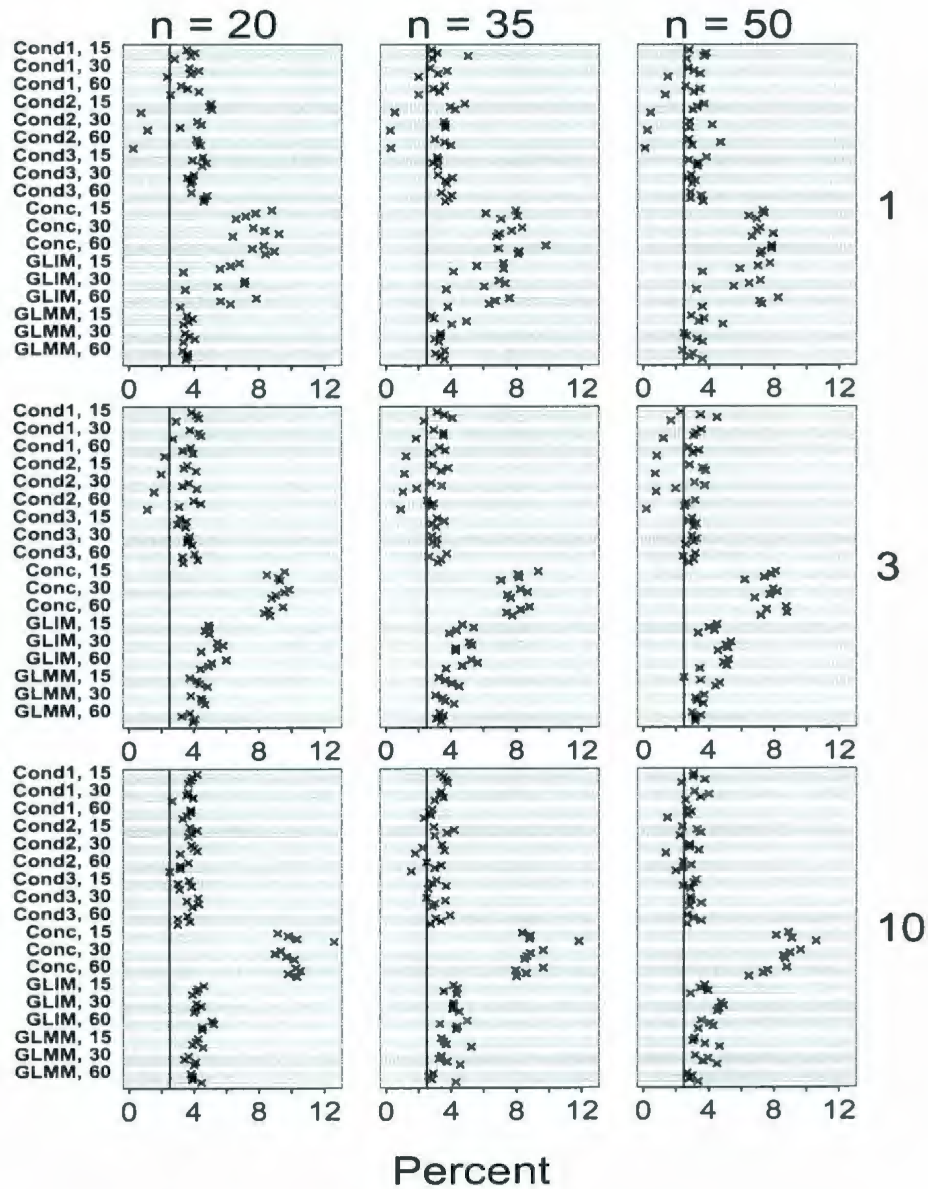

Figure D.33: Simulated lower exceedance of 95% confidence intervals. See Figure D.27 for details. 2.5% is shown as a solid vertical line.

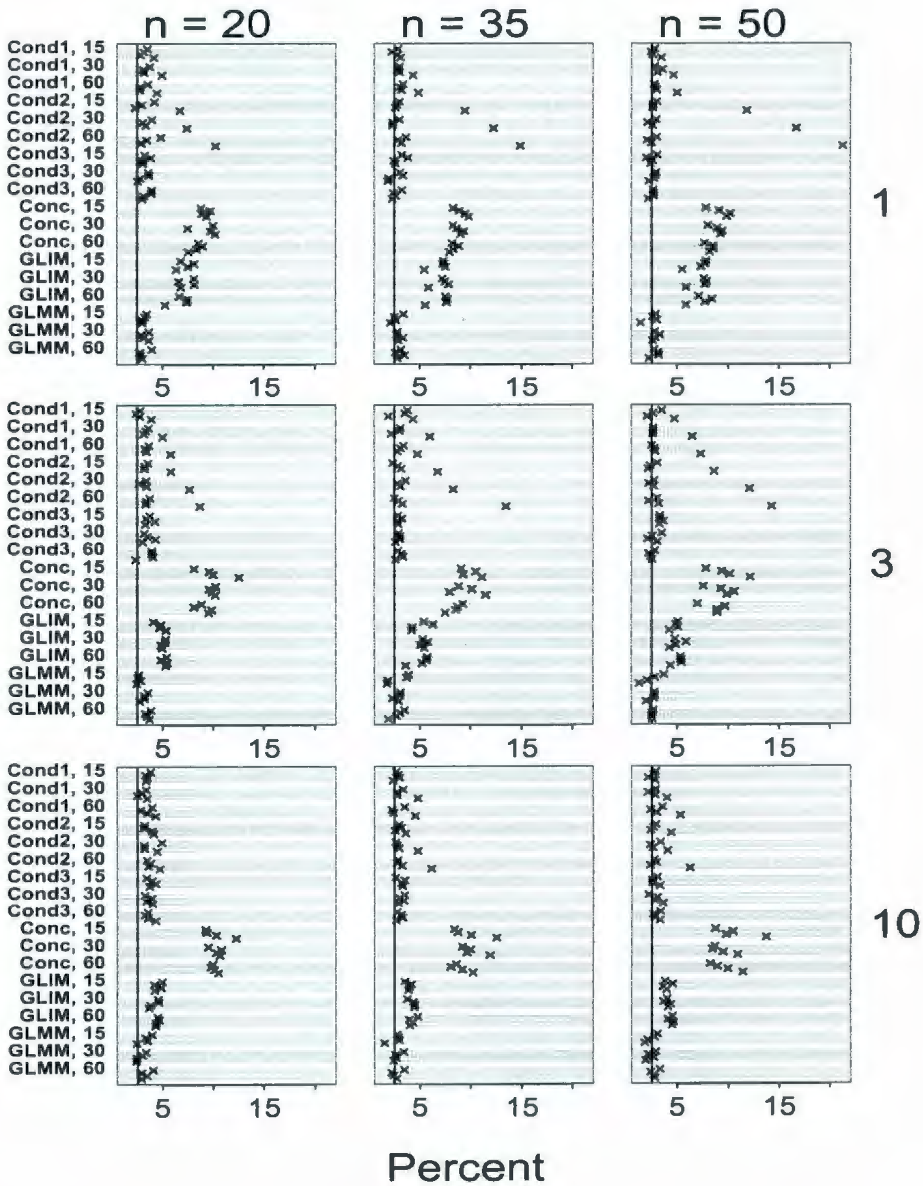

Figure D.34: Simulated upper exceedance of 95% confidence intervals. See Figure D.27 for details. 2.5% is shown as a solid vertical line.

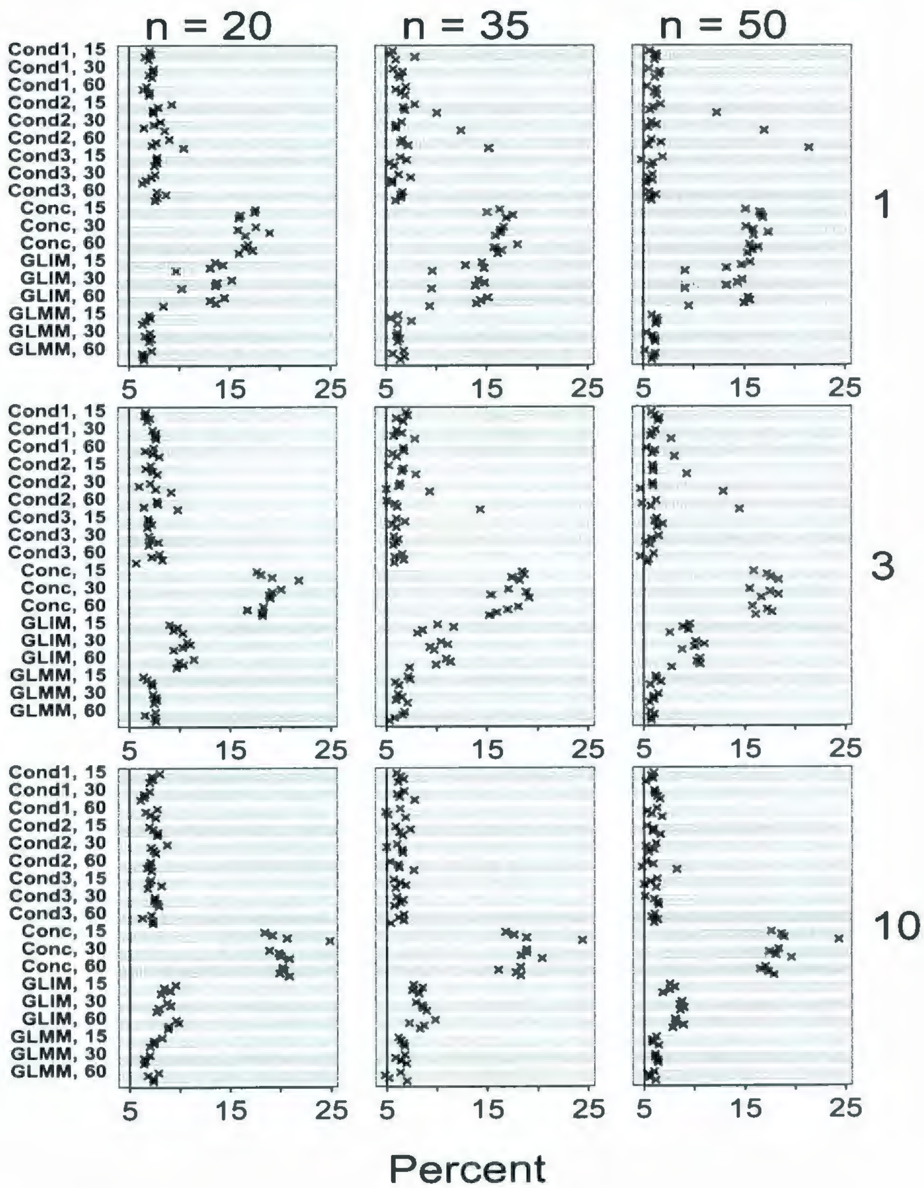

Figure D.35: Simulated total exceedance of 95% confidence intervals. See Figure D.27 for details. 5% is shown as a solid vertical line.

Bibliography

[1] Abramowitz, M. and Stegun, I.A. (1972) Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, U.S. Government Printing Office, Washington, D.C.

[2] Barndorff-Neilsen, O.E & Cox, D.R. (1994). Inferences and Asymptotics. London: Chapman and Hall.

[3] Benoît, H.P., and Swain, D.P. (2003). Accounting for length–and depth dependent diel variation in catchability of fish and invertebrates in an annual bottom–trawl survey. ICES J. Mar. Sci. **60**: 1298–1317.

[4] Cadigan, N.G., S.J. Walsh, and W. Brodie. (2006). Relative efficiency of the Wilfred Templeman and Alfred Needler research vessels using a Campelen 1800 shrimp trawl in NAFO Subdivision 3Ps and Divisions 3LN. DFO Canadian Science Advisory Secretariat Research Document 2006/085.

[5] Cameron, A.C. and Trivedi, P.K. (1998). Regression Analysis of Count Data. Econometric Society Monographs, vol. 30, Cambridge, Cambridge University Press.

[6] Collett, D. (1991) Modelling Binary Data. Chapman and Hall.

[7] Cox, D. R. (1970). The Analysis of Binary Data. London: Methuen.

[8] Cox, D.R., and Snell, E.J. (1989). Analysis of Binary Data. 2nd ed. Chapman and Hall, London.

[9] Davidson, R., and J.G. MacKinnon (1993), estimation and Inference in Econometrics, Oxford, Oxford University Press.

[10] Doubleday, W.G. [Ed.] 1981. Manual on groundfish surveys in the Northwest Atlantic. Northwest Atlantic Organization Scientific Council Studies, **2**: 7-55.

[11] Fanning, L.P. (1985). Intercalibration of research survey results obtained by different vessels. CAFSAC (Canadian Atlantic Fisheries Scientific Advisory Committee) Research Document **85/3**.

[12] Greenwood, M., and G.U. Yule (1920). "An inquiry into the Nature of Frequency Disributions of Multiple happenings, with Particular Reference to the Occurence of Multiple Attacks of Disease or Repeated Accidents,Journal of the Royal Statistical Society A, **83**, 255 279

[13] Grimmett, G. and Stirzaker, D. Probability and Random Processes, 2nd ed. Oxford, England: Oxford University Press, 1992.

[14] Gunderson, D. R. (1993). Surveys of Fisheries Resources. Wiley & Sons, New York.

[15] Hausman, J.A., B.H. Hall, and Z. Griliches (1984), "Econometric Models for Count Data with an application to the patents-R and D Relationships," Econometrica, **52**, 909 938.

[16] Heagerty, P.J., and Kurland, B.F. 2001. Misspecified maximum likelihood estimates and generalised linear mixed models. Biometrika, **88**: 973-985.

[17] James H. Power and E. Barry Moser. (1999). Linear model analysis of net catch data using the negative binomial distribution. Can. J. Fish. Aquat. Sci. **56**: 191 200.

[18] Lambert, D. (1992), "Zero-Inflated Poisson Regression with an Application to Defects in Manufacturing," Technometrics, **34**, 1–14.

[19] Lewy, P., Nielsen, J.R., and Hovgard, H. (2004). Survey gear calibration independent of apatial fish distribution. Can. J. Fish. Aquat. Sci. **61**: 636 647.

[20] McCullagh, P., and Nelder, J.A. (1989). Generalized linear models. 2nd ed. Chapman and Hall, London.

[21] Millar, R.B. (1992). Estimating the size selectivity of fishing gear by conditioning on the total catch. JASA **87**: 962 968.

[22] Mullahy, J. (1986), "Specification and Testing of Some Modified Count Data Models," Journal of Econometrics, **33**, 341–365.

[23] Pelletier, D. (1998). Intercalibration of research survey vessels in fisheries: a review and an applicatio. Can. J. Fish. Aquat. Sci. **55**: 2672 2690.

[24] Reid, N. (1995). The roles of conditioning in inference. Stat. Sci. **10**: 138 199.

[25] Smedbol, R.K., Shelton, P.A., Swain, D.P., Frechet, A., and Chouinard, G.A. 2002. Review of population structure, distribution and abundance of cod (Gadus morhua) in Atlantic Canada in a species-at-risk context. DFO Can. Sci. Advis. Sec. Res. Doc. 2002/082.

[26] Stefanski, L.A. (2000). Measurement error models. Journal of the American Statistical Association, **95**, 1353-1358.

[27] Stigler, S.M. (1986) The History of Statistics. Belknap Press, Cambridge, Mass.

[28] Williams, D.A. (1982) Extra-binomial variation in logistic linear models. Applied Statistics, **31**, 144–148.